INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

An elementary introduction to Malliavin calculus

Vlad Bally

N° 4718

Fevrier 2003

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4718--FR+ENG

An elementary introduction to Malliavin calculus

Vlad Bally

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Mathfi

Rapport de recherche n° 4718 — Fevrier 2003 — 49 pages

Abstract: We give an introduction to Malliavin calculus following the notes of four lectures that I gave in the working group of the research team Mathfi in October 2001. These notes contain three topics:

1. An elementary presentation of Malliavin's differential operators and of the integration by parts formula which represents the central tool in this theory.
2. The Wiener chaos decomposition and the differential operators of Malliavin calculus as operators on the Wiener space.
3. The application of this calculus to the study of the density of the law of a diffusion process. This was the initial application of Malliavin's calculus - and provides a probabilistic proof of Hormander's hypothesis.

The aim of these notes is to emphasis some basic ideas and technics related to this theory and not to provide a text book. The book of Nualart, for example, is an excellent nomography on this topic. So we live out some technical points (we send to the papers or books where complete proofs may be found) and do not treat problems in all generality or under the more general assumptions. We hope that this choice provides a soft presentation of the subject which permits to understand better the objects and the ideas coming on as well as the possible applications.

Key-words: Malliavin's Calculus, Wiener chaos decomposition, Integration by parts, Hormander's theorem.

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Une introduction élémentaire au Calcul de Malliavin

Résumé : On presente une introduction au calcul de Malliavin suivent les notes d'une suite de quatre éxposés que j'ai fait dans le cadre du groupe de travail de l'équipe Mathfi en octobre 2001. Ils contient trois sujets:

1. Une présentation élémentaire des opérateurs différentiels du calcul de Malliavin et de la formule d'intégration par partie qui est l'instrument central de cette théorie.

2. La décomposition en chaos de Wiener et la présentation des opérateurs différentiels sur cet espace.

3. L'application du Calcul de Malliavin à l'étude de la densité de la loi d'une diffusion. C'est l'application initiale de cette théorie qui a donné une démonstration probabiliste du théorème de Hormander.

Le but de ces notes n'est pas de donner une présentation complète de la théorie mais juste de mettre en évidence les objets pricipaux et les idée et techniques employés. Il y a dèja d'excellentes monografies sur le sujet, comme celle de Nualart par éxemple. On a donc laissé de côté certains points très techniques et on n'a pas forcement donné les resultats sous les hypothéses les plus générales. On espère que ceci a permit de donner une présentation allégée qui permet de mieux comprendre les objets et les idées ansi que les possibles applications.

Mots-clés : Calcul de Malliavin, Décomposition en chaos de Wiener, Intégration par parties, Théorème de Hormander.

1 Introduction

Malliavin calculus was conceived in the years $70's$ and in the years $80's$ and $90's$ a huge amount of work has been done in this field. It becomes an analysis on the Wiener space and several monographs on this subject are available nowadays: Nualart [15], Oksendal [16], Ikeda Watanabe [10]. The main application of Malliavin calculus was to give sufficient conditions in order that the law of a random variable has a smooth density with respect to Lebegue's measure and to give bounds for this density and its derivatives. In his initial papers Malliavin used the absolute continuity criterion in order to prove that under Hormander's condition the law of a diffusion process has a smooth density and in this way he gave a probabilistic proof of Hormander's theorem. Afterwards people used this calculus in various situations related with Stochastic $PDE's$ and so on. These last years Malliavin calculus found new applications in probabilistic numerical methods, essentially in the field of mathematical finance. These applications are quit different from the previous ones because the integration by parts formula in Malliavin calculus is employed in order to produce some explicit weights which come on in non linear algorithms.

At list considered from the point of view of people working in concrete applied mathematics, Malliaivn calculus appears as a rather sophisticated and technical theory which requires an important investments, and this may be discouraging.. The aim of these notes is to give an elementary introduction to this topic. Since the priority is to the easy access we allow ourself to be rather informal on some points, to avoid too technical proofs, to live out some developments of the theory which are too heavy and, at list for the moment, are not directly used in applications. We try to set up a minimal text which gives an idea about the basic objects and issues in Malliaivn calculus in order to help the reader to understand the new literature concerning applications on one hand and to accede to much more complete texts (as the monographs quoted above) on the other hand. The paper is organized as follows. In Section 1 we give an abstract integration by parts formula and investigate it's consequences for the study of the density of the law of a random variable, computation of conditional expectations and sensitivity with respect to a parameter. These are essentially standard distribution theory reasonings in a probabilistic frame. In Section 2 we present Malliavin's calculus itself. We start with simple functionals which are random variables of the form $F = f(\Delta_n^0, ..., \Delta_n^{2^n-1})$ where $\Delta_n^i = B(\frac{i+1}{2^n}) - B(\frac{i}{2^n})$ and f is a smooth function. So a simple functional is just an aggregate of increments of the Brownian motion B and this is coherent with probabilistic numerical methods - think for instance to the Euler scheme for a diffusion process. It turns out that at the level of simple functionals Malliavin's differential operators are qui similar to usual differential operators of R^{2^n}. The spacial fact which gives access to an infinite dimensional calculus is that $\Delta_n^i = \Delta_{n+1}^{2i} + \Delta_{n+1}^{2i+1}$. This permits to embed $C^\infty(R^{2^n}; R)$ into $C^\infty(R^{2^{n+1}}; R)$. So one obtains a chain of finite dimensional spaces, and Malliavin's derivative operators, which are defined on the finite dimensional spaces, are independent of the dimension. This allows to extend these operators (using L^2 norms) on infinite dimensional spaces. A key part in the extension procedure is played by the duality between the Malliavin derivative D and the Skorohod integral δ -because the duality relation guaranties that the two operators are closable. Next we discuss the relation between the differential operators and the chaos expansion on the Wiener space. It turns out that they have a very nice and explicit expression on the Weiner chaoses and this permits to precise their domains in terms of the speed of convergence of the chaos expansion series.. Finally we prove Malliavin's integration by parts formula and give the coresponding absolute continuity criterion. In Section 3 we give the applications of Malliavin calculus to diffusion process.

2 Abstract integration by parts formula

The central tool in the Malliavin calculus is an integration by parts formula. This is why we give in this section an abstract approach to the integration by parts formula and to the applications of this formula to the study of the density of the law of a random variable and to computation of conditional expectations. All this represent standard facts in the distribution theory but we present them in a probabilistic language which is appropriate for our aims. For simplicity reasons we present first the one dimensional case and then we give the extension to the multi-dimensional frame.

2.1 The one dimensional case

The density of the law Let (Ω, F, P) be a probability space and let $F, G : \Omega \to R$ be integrable random variables. We say that the integration by parts formula $IP(F;G)$ holds true if there exists an integrable random variable $H(F;G)$ such that

$$IP(F;G) \quad E(\phi^{'}(F)G) = E(\phi(F)H(F;G)), \quad .\forall \phi \in C_c^{\infty}(R) \tag{1}$$

where $C_c^{\infty}(R)$ is the space of infinitely differentiable functions with compact support.

Moreover, we say that the integration by parts formula $IP_k(F;G)$ holds true if there exists an integrable random variable $H_k(F;G)$ such that

$$IP_k(F;G) \quad E(\phi^{^{(k)}}(F)G) = E(\phi(F)H_k(F;G)), \quad .\forall \phi \in C_c^{\infty}(R) \tag{2}$$

Note that $IP(F;G)$ coincides with $IP_1(F;G)$ and $H(f;G) = H_1(F;G)$. Moreover, if $IP(F;G)$ and $IP(F;H(F;G))$ hold true then $IP_2(F;G)$ holds true with $H_2(F;G) = H(F;H(F;G))$. An analogues assertion holds true for higher order derivatives. This lids us to define $H_k(1)$ by recurrence: $H_0(F) = 1, H_{k+1}(F) = H(F;H_k(F))$.

The weight $H(F;G)$ in $IP(F;G)$ is not unique: for any R such that $E(\phi(F)R) = 0$ one may use $H(F;G) + R$ as well. In numerical methods this plays an important part because, if on wants to compute $E(\phi(F)H(F;G))$ using a Monte Carlo method then one would like to work with a weight which gives minimal variance (see $[F\S all, 1]$ and $[F\S all, 2]$). Note also that in order to perform a Monte Carlo algorithm one has to simulate F and $H(F;G)$. In some particular cases $H(F;G)$ may be computed directly, using some ad-hoc methods. But Malliavin's calculus gives a systhematic access to the computation of this weight. Typically in applications F is the solution of some stochastic equation and $H(F;G)$ appears as an aggregate of differential operators (in Malliavin's sense) acting on F. These quantities are also related to some stochastic equations and so one may use some approximations of these equations in order to produce concrete algorithms.

Let us give a simple example. Take $F = f(\Delta)$ and $G = g(\Delta)$ where f, g are some differentiable functions and Δ is a centered gaussian random variable of variance σ. Then

$$E(f'(\Delta)g(\Delta)) = E(f(\Delta)(g(\Delta)\frac{\Delta}{\sigma} - g'(\Delta))) \tag{3}$$

so $IP(F;G)$ holds true with $H(F;G) = g(\Delta)\frac{\Delta}{\sigma} - g'(\Delta)$.

The proof is a direct application of the standard integration by parts, but in the presence of the gaussian density $p(x) = \frac{1}{\sqrt{2\pi\delta}}\exp(-\frac{x^2}{2\sigma})$:

$$\begin{aligned} E(f'(\Delta)g(\Delta)) &= \int f'(x)g(x)p(x)dx = -\int f(x)(g'(x)p(x) + g(x)p'(x))dx \\ &= -\int f(x)(g'(x) + g(x)\frac{p'(x)}{p(x)})p(x)dx = E(f(\Delta)(g(\Delta)\frac{\Delta}{\sigma} - g'(\Delta))) \end{aligned}$$

Malliavin calculus produces the weights $H(F;G)$ for a large class of random variables - (3) represents the simplest example of this quind - but this is not the subject of this section. Here we give some consequences of the above property.

Lemma 1 *Suppose that F satisfies $IP(F;1)$. Then the law of F is absolutely continuous with respect to the Lebesgue measure and the density of the law is given by*

$$p(x) = E(1_{[x,\infty)}(F)H(F;1)). \tag{4}$$

Proof. The formal argument is the following: since $\delta_0(y) = \partial_y 1_{[0,\infty)}(y)$ one employes $IP(F;1)$ in order to obtain

$$E(\delta_0(F-x)) = E(\partial_y 1_{[0,\infty)}(F-x)) = E(1_{[0,\infty)}(F-x)H_1(F;1)) = E(1_{[x,\infty)}(F)H(F;1)).$$

In order to let this reasoning rigorous one has to regularize the Dirac function. So we take a positive function $\phi \in C_c^\infty(R)$ with the support equal to $[-1,1]$ and such that $\int \phi(y)dy = 1$ and for each $\delta > 0$ we define $\phi_\delta(y) = \delta^{-1}\phi(y\delta^{-1})$. Moreover we define Φ_δ to be the primitive of ϕ_δ given by $\Phi_\delta(y) = \int_{-\infty}^y \phi_\delta(z)dz$ and we construct some random variables θ_δ of law $\phi_\delta(y)dy$ and which are independent of F. For each $f \in C_c^\infty(R)$ we have

$$E(f(F)) = \lim_{\delta\to 0} Ef(F-\theta_\delta)). \tag{5}$$

We compute now

$$\begin{aligned} Ef(F-\theta_\delta)) &= \int\int f(u-v)\phi_\delta(v)dvdP\circ F^{-1}(u) = \int\int f(z)\phi_\delta(u-z)dzdP\circ F^{-1}(u) \\ &= \int f(z)E(\phi_\delta(F-z))dz = \int f(z)E(\Phi'_\delta(F-z))dz \\ &= \int f(z)E(\Phi_\delta(F-z)H(F;1))dz. \end{aligned}$$

The above relation togather with (5) guarantees that the law of F is absolutely continuous with respect to the Lebesgue measure. On the other hand $\Phi_\delta(y) \to 1_{[x,\infty)}(y)$ except for $y=0$, so $\Phi_\delta(F-z) \to 1_{[0,\infty)}(F-z) = 1_{[z,\infty)}(F), P-a.s.$ Then using Lebesgues dominated convergence theorem we pass to the limit in the above relation and we obtain

$$E(f(F)) = \int f(z) E(1_{[z,\infty)}(F) H(F;1)) dz.$$

The above integral representation theorem gives some more information on the density p :

$\diamondsuit$ **Regularity**. Using Lebesgue's dominated convergence theorem one may easily check that $x \to p(x) = E(1_{[x,\infty)}(F)H(F;1))$ is a continuous function. We shall see in the sequel that, if one has some more integration by parts formulas then one may prove that the density is differentiable.

$\diamondsuit$ **Bounds** Suppose that $H(F;1)$ is square integrable. Then, using Chebishev's inequality

$$p(x) \leq \sqrt{P(F \geq x)} \left\| H(F;1) \right\|_2.$$

In particular $\lim_{x\to\infty} p(x) = 0$ and the convergence rate is controlled by the tails of the law of F. For example if F has finite moments of order p this gives $p(x) \leq \frac{C}{x^{p/2}}$. In significant examples, as diffusion processes, the tails have even exponential rate. So the problem of the upper bounds for the density is rather simple (at the contrary, the problem of lower bounds is much more challenging). The above formula gives a control for $x \to \infty$. In order to obtain similar bounds for $x \to -\infty$ one has to employ the formula

$$p(x) = E(1_{(-\infty,x)}(F)H(F;1)). \tag{6}$$

This is obtained in the same way as (4) using the primitive $\Phi_\delta(y) = -\int_y^\infty \phi_\delta(z)dz$.

We go now further and treat the problem of the derivatives of the density function.

Lemma 2 *Suppose that $IP(F; H_i(F)), i = 0, ..., k$ holds true. Then the density p is k times differentiable and*

$$p^{(i)}(x) = (-1)^i E(1_{(x,\infty)}(F) H_{i+1}(F)) \tag{7}$$

Proof. Let $i=1$. We define $\Psi_\delta(x) = \int_{-\infty}^x \Phi_\delta(y)dy$, so that $\Psi_\delta'' = \phi_\delta$, and we come back to the proof of lemma 1. We use twice the integration by parts formula and obtain

$$E(\phi_\delta(F-z)) = E(\Phi_\delta(F-z)H(F;1)) = E(\Psi_\delta(F-z)H(F;H(F;1))).$$

Since $\lim_{\delta\to 0} \Psi_\delta(F-z) = (F-z)_+$ we obtain

$$E(f(F)) = \int f(z) E((F-z)_+ H(F;H(F;1))) dz$$

and so

$$p(z) = E((F-z)_+ H(F; H(F;1))).$$

The pleasant point in this new integral representation of the density is that $z \to (F-z)_+$ is differentiable. Taking derivatives in the above formula gives

$$p'(z) = -E(1_{[z,\infty)}(F) H(F; H(F;1))) = -E(1_{[z,\infty)}(F) H_2(F))$$

and the proof is completed for $i = 1$. In order to deal with higher derivatives one employes more integration by parts in order to obtain

$$p(z) = E(\theta_i (F-z)_+ H_{i+1}(F))$$

where θ_i is an i times differentiable function such that $\theta_i^{(i)}(x) = 1_{[0,\infty)}(x)$. $\square$

Let us make some remarks.

$\diamondsuit$ Watanabe develops in [20] (see also [10]) a distribution theory on the Wiener space related to Malliavin calculus. One of the central ideas in this theory is the so called procedure of "pushing back Schwartz distributions" which essentially coincides with the trick used in our proof: one employes a certain number of integration by parts in order to regularize the distributions, which coresponds in our proof to the representation by means of the function θ_i which is i times differentiable.

$\diamondsuit$ The integral representation formula (7) permits to obtain upper bounds of the derivatives of the density p. In particular, suppose that F has finite moments of any order and that $IP(F; H_i(F))$ holds true for every $i \in N$ and $H_i(F)$ are square integrable. Then p is infinitely differentiable and $\left|p^{(i)}(x)\right| \leq \sqrt{P(F > x)}\, \|H_i(F)\|_2 \leq \frac{C}{x^{p/2}}$ for every $p \in N$. So $p \in \mathcal{S}$, the Schwartz space of rapidly decreasing functions.

$\diamondsuit$ The above lemma shows that there is an intimate relation (quasi equivalence) between the integration by parts formula and the existence of a "good" density of the law of F. In fact, suppose that $F \sim p(x)dx$ and p is differentiable and $p'(F)$ is integrable. Then, for every $f \in C_c^\infty(R)$

$$\begin{aligned} E(f'(F)) &= \int f'(x)p(x)dx = -\int f(x)p'(x)dx = -\int f(x)\frac{p'(x)}{p(x)} 1_{(p>0)}(x)p(x)dx \\ &= -E(f(F)\frac{p'(F)}{p(F)} 1_{(p>0)}(F)). \end{aligned}$$

So $IP(F, 1)$ holds with $H(F;1) = -\frac{p'(F)}{p(F)} 1_{(p>0)}(F) \in L^1$ (because $p'(F) \in L^1(\Omega)$). Iterating this we obtain the following chain of implications: $IP(F, H_{k+1}(F))$ holds true $\Rightarrow p$ is k times differentiable and $p^{(k)}(F) \in L^1(\Omega) \Rightarrow IP(F, H_{k-1}(F))$ holds true. Moreover one has $H_k(F) = (-1)^k \frac{p^{(k)}(F)}{p(F)} 1_{(p>0)}(F) \in L^1$.

Conditional expectations The computation of conditional expectations is crucial for solving numerically certain non linear problems by descending a Dynamical Programing algorithm. Recently several authors (see $[F\S al, 1l], [F\S all2], ...$) employed some formulas based on Malliavin calculus technics in order to compute conditional expectations. In this section we give the abstract form of this formula.

Lemma 3 *Let F and G be two real random variables such that $IP(F;1)$ and $IP(F;G)$ hold true. Then*

$$E(G \mid F = x) = \frac{E(1_{[x,\infty)}(F)H(F;G))}{E(1_{[x,\infty)}(F)H(F;1))} \tag{8}$$

with the convention that the term in the right hand side is null when the denominator is null.

Proof. Let $\theta(x)$ designees the term in the right hand side of the above equality. We have to check that for every $f \in C_c^\infty(R)$ one has $E(f(F)G) = E(f(F)\theta(F))$. Using the regularization functions from the proof of Lemma 1 we write

$$\begin{aligned} E(f(F)G) &= E(G \lim_{\delta\to 0} \int f(z)\phi_\delta(F-z)dz) = \lim_{\delta\to 0} \int f(z)E(G\phi_\delta(F-z))dz \\ &= \lim_{\delta\to 0} \int f(z)E(\Phi_\delta(F-z)H(F;G))dz = \int f(z)E(1_{[0,\infty)}(F-z)H(F;G))dz \\ &= \int f(z)\theta(z)p(z)dz = E(f(F)\theta(F)) \end{aligned}$$

and the proof is completed. □

The sensitivity problem In many applications one considers quantities of the form $E(\phi(F^x))$ where F^x is a family of random variables indexed on a finite dimensional parameter x. A typical example is $F^x = X_t^x$ which is a diffusion process starting from x. In order to study the sensitivity of this quantity with respect to the parameter x one has to prove that $x \to E\phi(F^x)$ is differentiable and to evaluate the derivative. There are two ways to take this problem: using a pathways approach or an approach in law.

The pathways approach supposes that $x \to F^x(\omega)$ is differentiable for almost every ω (and this is the case for $x \to X_t^x(\omega)$ for example) and ϕ is differentiable also. Then $\partial_x E(\phi(F^x)) = E(\phi'(F^x)\partial_x F^x)$. But this approach brakes down if ϕ is not differentiable. The second approach overcomes this difficulty using the smoothness of the density of the law of F^x. So, in this approach one assumes that $F^x \sim p^x(y)dy$ and $x \to p^x(y)$ is differentiable for each y. Then $\partial_x E(\phi(F^x)) = \int \phi(y)\partial_x p^x(y)dy = \int \phi(y)\partial_x \ln p^x(y)p^x(y)dy = E(\phi(F^x)\partial_x \ln p^x(F))$. Some times engineers call $\partial_x \ln p^x(F)$ the score function (see Vasquez-Abad [19]). But of course this approach works when one knows the density of the law of F^x. The integration by parts formula $IP(F^x, \partial_x F^x)$ permits to write down the equality

$\partial_x E(\phi(F^x)) = E(\phi'(F^x)\partial_x F^x) = E(\phi(F^x)H(F^x, \partial_x F^x))$ without having to know the density of the law of F^x (the above equality holds true even if ϕ is not derivable because there are no derivatives in the first and last term - so one may use some regularization and then pass to the limit). So the quantity of interest is the weight $H(F^x, \partial_x F^x)$. Malliavin calculus is a machinery which permits to compute such quantities for a large class of random variables for which the density of the law is not known explicitly (for example diffusion process). This is the approach in $[F\S all, 1]$ and $[F\S all, 2]$ to the computation of Greeks (sensitivity of the price of European and American options with respect to certain parameters) in mathematical finance problems.

2.2 The multi dimensional case

In this section we deal with a d - dimensional random variable $F = (F^1, ..., F^d)$ instead of the one dimensional random variable considered in the previous section. The results concerning the density of the law and the conditional expectation are quite analogous.. Let us introduce some notation. For a multi-index $\alpha = (\alpha_1, ..., \alpha_d) \in N^d$ we denote $|\alpha| = \sum_{i=1}^d \alpha_i$ and $D^\alpha = \partial^{|\alpha|}/\partial^{\alpha_1}...\partial^{\alpha_d}$, with the convention that ∂^0 is just the identity. The integration by parts formula is now the following. We say that the integration by parts formula $IP_\alpha(F;G)$ holds true if there exists an integrable random variable $H_\alpha(F;G)$ such that

$$IP_\alpha(F;G) \quad E(D^\alpha \phi(F)G) = E(\phi(F)H_\alpha(F;G)), \quad .\forall \phi \in C_c^\infty(R^d). \tag{9}$$

Let us give a simple example which turns out to be central in Malliavin calculus. Take $F = f(\Delta^1, ..., \Delta^m)$ and $G = g(\Delta^1, ..., \Delta^m)$ where f, g are some differentiable functions and $\Delta^1, ..., \Delta^m$ are independent, centered gaussian random variables with variances $\sigma^1, ..., \sigma^m$. We denote $\Delta = (\Delta^1, ..., \Delta^m)$. Then for each $i = 1, ..., m$

$$E(\frac{\partial f}{\partial x^i}(\Delta)g(\Delta)) = E(f(\Delta)(g(\Delta)\frac{\Delta^i}{\sigma^i} - \frac{\partial g}{\partial x^i}(\Delta))) \tag{10}$$

This is an immediate consequence of (3) and of the independence of $\Delta^1, ..., \Delta^m$. We give now the result concerning the density of the law of F.

Proposition 1 *i) Suppose that $IP_{(1,...,1)}(F;1)$ holds true. Then $F \sim p(x)dx$ with*

$$p(x) = E(1_{I(x)}(F)H_{(1,...,1)}(F;1)) \tag{11}$$

where $I(x) = \prod_{i=1}^d [x^i, \infty)$. In particular p is continuous.

ii) Suppose that for every multi-index $\alpha, IP_\alpha(F;1)$ holds true . Then $D^\alpha p$ exists and is given by

$$D^\alpha p(x) = (-1)^{|\alpha|} E(1_{I(x)}(F)H_{(\alpha+1)}(F;1)) \tag{12}$$

where $(\alpha + 1) =: (\alpha_1 + 1, ..., \alpha_d + 1)$. Moreover, if $H_\alpha(F;1) \in L^2(\Omega)$ and F has finite moments of any order then $p \in S$, the Schwartz space of infinitely differentiable functions which decrease rapidly to infinity, together with all their derivatives.

Proof. The formal argument for i) is based on $\delta_0(y) = D^{(1,\ldots,1)}1_{I(0)}(y)$ and the integration by parts formula. In order to let it rigorous one has to regularize the Dirac function as in the proof of Lemma 1. In order to prove ii) one employes the same "pushing back Scwartz distribution" argument as in the proof of Lemma 2. Finally, in order to obtain bounds we write

$$|D^\alpha p(x)| \leq \sqrt{P(F^1 > x^1, ..., F^d > x^d)} \left\| H_{(\alpha+1)}(F;1) \right\|_2 .$$

If $x^1 > 0, ..., x^d > 0$ this together with Chebishev's inequality yields $|D^\alpha p(x)| \leq C_q |x|^{-q}$ for every $q \in N$. If the coordinates of x are not positive we have to use a variant of (12) which involves $(-\infty, x^i]$ instead of (x^i, ∞). □

The result concerning the conditional expectation reads as follows.

Proposition 2 *Let F and G be two real random variables such that $IP_{(1,\ldots,1)}(F;1)$ and $IP_{(1,\ldots,1)}(F;G)$ hold true. Then*

$$E(G \mid F = x) = \frac{E(1_{I(x)}(F)H(F;G))}{E(1_{I(x)}(F)H(F;1))} \tag{13}$$

with the convention that the term in the right hand side is null when the denominator is null.

Proof. The proof is the same as for Lemma 3 but one has to use the regularization function $\overline{\phi}_\delta(x) = \prod_{i=1}^d \phi_\delta(x^i)$ and $\overline{\Phi}_\delta(x) = \prod_{i=1}^d \Phi_\delta(x^i)$ and the fact that $D^{(1,\ldots,1)}\overline{\Phi}_\delta(x) = \overline{\phi}_\delta(x)$. □

3 Malliavin Calculus

3.1 The differential and the integral operators in Malliavin's Calculus

Let (Ω, F, P) be a probability space, $B = (B_t)_{t\geq 0}$ a $d-$ dimensional Brownian motion and $(F_t)_{t\geq 0}$ the filtration generated by B, that is: F_t is the completation of $\sigma(B_s, s \leq t)$ with the $P-$null sets. The fact that we work with the filtration generated by the Brownian motion and not with some abstract filtration is central: all the objects involved in Malliavin calculus are 'functionals' of the Brownian motion and so we work in the universe generated by B. The first step will be to consider 'simple functionals', *i.e.* smooth functions of a finite number of increments of the Brownian motion. We set up an integro-differential calculus on these functionals, which is in fact a classical finite dimensional calculus. The cleaver thing is that the operators in this calculus does not depend on the dimension (the number of increments involved in the simple functional) and so they admit an infinite dimensional extension.. Another important point is that an integration by parts formula holds true. On one hand this formula permits to prove that the operators defined on simple functionals are closable - which is necessary in order to get an extension. On the other hand the integration by parts formula itself extends to general functional and this represents one of the main tools in Malliavin calculus (the previous section gives sufficient motivation for such a formula).

Simple functionals and simple processes Let us introduce some notation. We put $t_n^k = \frac{k}{2^n}$ and $I_n^k = [t_n^k, t_n^{k+1})$. Note that $I_n^k = I_{n+1}^{2k} \cup I_{n+1}^{2k+1}$. This simple fact is important in order to define operators which does not depend on the dimension of the space on which we work. Moreover we denote $\Delta_n^k = B(t_n^{k+1}) - B(t_n^k)$. On components this reads $\Delta_n^k = (\Delta_n^{k,1}, ..., \Delta_n^{k,d})$ with $\Delta_n^{k,i} = B^i(t_n^{k+1}) - B^i(t_n^k)$. We define the simple functionals of order n to be the random variables of the form

$$F = f(\Delta_n^0, ..., \Delta_n^{m-1}) \tag{14}$$

where $m \in N$ and $f : R^{m\times d} \to R$ is a smooth function with polynomial growth. In particular F has finite moments of any order. We denote by S_n the space of all the simple functionals of order n and $S = \cup_{n\in N} S_n$ is the space of all the simple functionals. We have the following fundamental fact (which we do not prove here)

Proposition 3 *For every $p > 0$, S is a linear subspace of $L^p(\Omega, F_\infty, P)$ and it is dense.*

One may choose any other class of functions f in the definition of the simple functionals (for example polynomials or smooth functions with compact support) but has to take care that the density property holds true. The fact that the filtration is generated by the Brownian motion is the reason for which we have the density property.

The representation (14) has the following uniqueness property: if $F \in S_n$ and $F = f(\Delta_n^0, ..., \Delta_n^{m-1}) = g(\Delta_n^0, ..., \Delta_n^{m'-1})$ then $f = g$ and $m = m'$. So we may denote by f_n the unique function which represents F as a simple functional of order n. On the other hand if $F \in S_n$ then $F \in S_{n'}$ for every $n' > n$ and the representation functions f_n and $f_{n'}$ are related by the following recursive relation:

$$\begin{aligned} f_{n+1}(x) &= f_n(s_{n+1}(x)), \quad with & (15)\\ s_{n+1}(x^0, x^1, ..., x^{2i}, x^{2i+1}, ..., x^{2m-2}, x^{2m-1}) &= (x^0 + x^1, ..., x^{2i} + x^{2i+1}, ..., x^{2m-2} + x^{2m-1}) \end{aligned}$$

The above property will permit to define the differential operators on the space S_n and then to check that this definition does not depend on n and so to obtain an operator on S. Moreover, since S is dense in $L^2(\Omega, F_\infty, P)$ we may define the extension of the differential operators (after proving that they are closable).

We keep m to be a free positive integer because we deal with the calculus in infinite time horizonte, but, if we want to construct Malliavin calculus in finite time horizonte, say $t \in [0, 1]$, then for each fixed n one has to take $m = 2^n$ which is the number of intervals I_n^k which cover the whole time interval $[0, 1]$. Let us consider for a moment this simpler situation. Then a simple functional of order n is of the form $F = f(\Delta_n^0, ..., \Delta_n^{2^n-1})$. Let us denote $C_n^\infty = C_p^\infty(R^{2^n}; R)$, the space of infinitely differentiable functions which has polynomial growth. Then the uniqueness of the representation give a bijection $i_n : C_n^\infty \to S_n$. On the other hand, by means of the functions s_n we construct the embedding $j_n : C_n^\infty \to C_{n+1}^\infty$ given by $j_n(f) = f \circ s_n$ and the corresponding embedding $q_n : S_n \to S_{n+1}$. So we have the following chains of embedding

$$\begin{aligned} S_1 &\subset \quad ...S_n \subset S_{n+1} \subset ... \subset S = \bigcup_{n\in N} S_n \\ C_1^\infty &\subset \quad ...C_n^\infty \subset C_{n+1}^\infty \subset ... \subset C_\infty^\infty =: \bigcup_{n\in N} C_n^\infty . \end{aligned} \tag{16}$$

We define now the simple processes of order n to be the processes of the form

$$u(t,\omega) = \sum_{i=0}^{m-1} u_k(\Delta_n^0, ..., \Delta_n^{m-1}) 1_{I_n^k}(t) \tag{17}$$

where $u_k : R^{m\times d} \to R, k = 0, ..., m-1$ are smooth functions with polynomial growth. So $u_k(\Delta_n^0, ..., \Delta_n^{m-1})$ is a simple functional of order n. Note that u_k depends on all the increments so the process u is not adapted. We denote by P_n the class of all the simple processes of order n and by P the class of all the simple processes $P = \cup_{n\in N} P_n$. We have

Proposition 4 *P is a dense subspace of $L^p([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ where B_+ is are the Borel sets of R_+ and λ is the Lebesgue measure.*

The functions u_k in the above representation are unique so we denote by u_k^n the functions in the representation of order n. Clearly a simple process of order n is also a simple process of order $m > n$ and we have the recurrence formulas

$$u_{2k}^{n+1}(x) = u_{2k+1}^{n+1}(x) = u_k^n(s_{n+1}(x)). \tag{18}$$

Let us come back to the finite time interval *i.e.* $t \in [0,1]$ and $m = 2^n$ (we use the same notation as above).We denote $C_{n,n}^\infty = C_p^\infty(R^{2^n}; R^{2^n})$, the space of $R^{2^n}-$valued infinitely differentiable functions with polynomial growth and, for $u \in C_{n,n}^\infty$ we denote by u_k the $k'th$ component. Then the uniqueness of the representation gives a bijection $i_{n,n} : C_n^\infty \to P_n$. On the other hand, by means of the functions s_n we construct the embedding $j_{n,n} : C_{n,n}^\infty \to C_{n+1,n+1}^\infty$ given by $j_{n+1,n+1}(u)_{2k} = j_{n+1,n+1}(u)_{2k+1} = u_k \circ s_{n+1}$ and the corresponding embedding $q_n : P_n \to P_{n+1}$. So we have the following chains of embedding

$$\begin{aligned} P_1 &\subset \quad ...P_n \subset P_{n+1} \subset ... \subset P = \bigcup_{n\in N} P_n \\ C_1^\infty &\subset \quad ...C_{n,n}^\infty \subset C_{n+1,n+1}^\infty \subset ... \subset C_{\infty,\infty}^\infty =: \bigcup_{n\in N} C_{n,n}^\infty . \end{aligned} \tag{19}$$

3.1.1 Malliavin derivatives and Skorohod integral.

We define the Malliavin derivative $D^i : S \to P, i = 1, ..., d$ by

$$D_t^i F = \sum_{k=1}^{m-1} \frac{\partial f_n}{\partial x^{k,i}} (\Delta_n^0, ..., \Delta_n^{m-1}) 1_{I_n^k}(t). \tag{20}$$

Let us precise the notation. We have $x = (x^0, ..., x^{m-1})$ with $x^k = (x^{k,1}, ..., x^{k,d})$. So x^k corresponds to the increment Δ_n^k and the $i'th$ component $x^{k,i}$ corresponds to the $i'th$ component $\Delta_n^{k,i}$. It is easy to check that the above definition does not depend on n and consequently is correct. From an intuitive point of view $D_t^i F$ represents the derivative of F with respect to the increment of B^i coresponding to t. We will sometimes use the following intuitive notation

$$D_t^i F = \frac{\partial F}{\partial \Delta_t^i}$$

where $\Delta_t^i = B^i(t_n^{k+1}) - B^i(t_n^k) = \Delta_n^{k,i}$ for $t_n^k \le t < t_n^{k+1}$.

We define the Skorohod integral $\delta^i : P \to S, i = 1, ..., d$ by

$$\delta^i(u) = \sum_{k=1}^{m-1} u_n^k(\Delta_n^0, ..., \Delta_n^{m-1}))\Delta_n^{k,i} - \sum_{k=1}^{m-1} \frac{\partial u_k^n}{\partial x^{k,i}} (\Delta_n^0, ..., \Delta_n^{m-1}) \frac{1}{2^n}. \tag{21}$$

and note that the definition does not depend on n and so is correct.

Note also that if u is a previsible process (*i.e.* u_k^n depends on $\Delta_n^0, ..., \Delta_n^{k-1}$ only) then $\frac{\partial u_k^n}{\partial x^{k,i}} = 0$ and so $\delta^i(u)$ coincides with the standard Ito integral with respect to B^i.

A central fact is that the operators D^i and δ^i are adjoint:

Proposition 5 *For every $F \in S$ and $u \in P$*

$$E \int_0^\infty D_t^i F \times u(t, \omega) dt = E(F \delta^i(u)). \tag{22}$$

Proof. We take n sufficiently large in order to have $F \in S_n$ and $u \in P_n$ and we take f_n, respectively $u_n^k, k = 0, ..., m-1$, the functions which give the representation of F, respectively of u (it is clear that one may choose the same m in the representation of F and u as well). We also denote $\Delta_n = (\Delta_n^0, ..., \Delta_n^{m-1})$. As a consequence of the integration by parts formula (10)

$$\begin{aligned} E \int_0^1 D_t^i F \times u(t, \omega) dt &= E \sum_{k=0}^{m-1} \frac{\partial f_n}{\partial x^{k,i}} (\Delta_n) u_n^k(\Delta_n) \frac{1}{2^n} \\ &= E(f_n(\Delta_n)(\sum_{k=0}^{m-1} u_n^k(\Delta_n)\Delta_n^{k,i} - \sum_{k=1}^{m-1} \frac{\partial u_k^n}{\partial x^{k,i}}(\Delta_n)\frac{1}{2^n})) = E(F\delta^i(u)). \end{aligned}$$

□

The above proposition gives the possibility to prove that D and δ are closable and so to define their extensions.

Proposition 6 *i)* $D^i : S \subset L^2(\Omega, F_\infty, P) \to P \subset L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ *is a linear closable operator. We define* $Dom(D^i)$ *to be the subspace of all the random variables* $F \in L^2(\Omega, F_\infty, P)$ *such that there exists a sequence of simple functionals* $F_n \in S, n \in N$ *such that* $F_n \to F$ *in* $L^2(\Omega, F_\infty, P)$ *and* $D^i F_n, n \in N$ *is a Cauchy sequence in* $L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ *We define* $D^i F =: \lim_n D^i F_n$.

ii) $\delta^i : P \subset L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P) \to S \subset L^2(\Omega, F_\infty, P)$ *is a linear closable operator. We define* $Dom(\delta^i)$ *to be the subspace of all the measurable processes* $u \in L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ *such that there exists a sequence of simple processes* $u_n \in P, n \in N$ *such that* $u_n \to u$ *in* $L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ *and* $\delta^i(u_n), n \in N$ *is a Cauchy sequence in* $L^2(\Omega, F_\infty, P)$ *We define* $\delta^i(u) =: \lim_n \delta(u_n)$.

iii) For every $F \in Dom(D^i)$ *and* $u \in Dom(\delta^i)$

$$E \int_0^1 D_t^i F \times u(t,\omega) dt = E(F \delta^i(u)). \tag{23}$$

Proof. Let $F_n \in S, n \in N$ such that $F_n \to 0$ in $L^2(\Omega, F_\infty, P)$ and such that $D^i F_n \to u$ in $L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$. In order to check that D^i is closable we have to prove that $u = 0$. This guarantees that the definition of D^i does not depend on the approximation sequencle F_n. Take $v \in P$. Using the duality (22)

$$E \int_0^1 u(t,\omega) v(t,\omega) dt = \lim_n E \int_0^1 D_t^i F_n v(t,\omega) dt = \lim_n E(F_n \delta^i(v)) = 0.$$

Since P is dense in $L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ it follows that $u = 0$. The fact that δ^i is closable is proved in the same way and (23) follows from (22) by approximation with simple functionals and simple processes. □

Remark 1 *Note that* $D_t^i F$ *is a stochastic process, but not an adapted one - it depends on the whole path of the Brownian motion. Note also that this process is defined as an element of* $L^2([0,\infty) \times \Omega, B_+ \times F_\infty, \lambda \times P)$ *so* $D_t^i F$ *is not determined for each fixed* t. *This lids sometimes to delicate problems but usually one is able to precise a privileged (for example continuous) version of it. In many situations of interest* F *appears as the solution of some stochastic equation and then* $D^i F$ *itself satisfies some equation obtained by "taking derivatives" in the equation of* F. *So the privileged version is defined as the solution of this new equation.*

Remark 2 *It is easy to check using approximation with simple functionals that, if* $F \in Dom(D^i)$ *is* F_t *measurable then* $D_s^i F = 0$ *for* $s > t$. *This corresponds to the intuitive idea that* D_s *represents the derivative with respect to the noise in* s *and sinc* F *does not depend on the Brownian motion* B_s *after* t *one has* $D_s^i F = 0$.

Remark 3 *If u is a previsible square integrable process then $u \in Dom(\delta^i)$ and $\delta^i(u)$ is the standard Ito stochastic integral with respect to B^i. This follows easily by approximation with simple functionals.*

Remark 4 *Most of the people working in Malliavin calculus use a slightly different definition of the simple functionals and of the Malliavin derivative. A simple functional is a random variable of the form $F = f(B(h^1), ..., B(h^m))$ where $h^k = (h^{k,1}, ..., h^{k,d}) \in L^2([0, \infty), B_+, \lambda)$ and $B^i(h^{k,i}) = \int_0^\infty h_s^{k,i} dB_s^i$. Then one defines*

$$D_t^i F = \sum_{k=1}^m \frac{\partial f}{\partial x^{k,i}} (B(h^1), ..., B(h^m)) h^{k,i}(t). \tag{24}$$

So the space of simple functionals is larger - note that $\Delta_n^{k,i} = B^i(h^{k,i})$ with $h^{k,i}(t) = 1_{[t_n^k, t_n^{k+1})}(t)$. But it is easy to check (using approximations of $h^{k,i}$ with step functions) that $F = f(B(h^1), ..., B(h^m)) \in Dom(D^i)$ (in the sense given here to $Dom(D^i)$) and $D_t^i F$ is given by the expression in the right hand side of (24). So the theory is the same. We prefer to use Δ_n^k as the basic elements because the simple functionals appear as aggregates of Brownian increments and this is the point of view in probabilistic numerical methods (think for example to the Euler scheme).

We define on S the norm

$$\begin{aligned} \|F\|_{1,2}^2 &= E\left|F\right|^2 + E < DF, DF > \quad \textit{with} \\ &< DF, DG >=: \int_0^\infty \sum_{i=1}^d D_t^i F D_t^i G dt \end{aligned}$$

and we define $D^{1,2}$ to be the closure of S with respect to this norm. Then $D^{1,2} = Dom(D) \subset L^2(\Omega, F_\infty, P)$.

Remark 5 *There is a strong analogy between the construction here and the standard Sobolev spaces in finite dimension. Let us see this in more detail. We consider the case when $t \in [0,1]$ so that $m = 2^n$ at level n, and we also assume that we have a single Brownian motion, that is $d = 1$. Let $n \in N$ be fixed and let $D \subset R^{2^n}$ be a bounded domain. For $f \in C_n^\infty = C_p^\infty(R^{2^n}; R)$ (we use the notation from the previous section) $\nabla f \in C_{n,n}^\infty = C_p^\infty(R^{2^n}; R^{2^n})$ and the Sobolev space $H^1(D) = W^{1,2}(D)$ is defined as the closure of C_n^∞ with respect to the norm $\|f\|_{D,1,2} = \|f\|_{L^2(\lambda_D)} + \sum_{i=1}^{2^n} \left\| \frac{\partial f}{\partial x^i} \right\|_{L^2(\lambda_D)}$ where $\lambda_D(dx) = 1_D(x) d\lambda(dx)$ with λ the Lebesgue measure. Recall now that in our frame C_n^∞ (respectively $C_{n,n}^\infty$) is in bijection with S_n (respectively with P_n) on which we consider the L^2 norm with respect to the probability P (respectively with respect to $\lambda \times P$). If we transport these measures by means of the isomorphisms we obtain the norm*

$$\|f\|_{n,1,2} = \|f\|_{L^2(\lambda_n)} + \sum_{i=1}^{2^n} \left\| \frac{\partial f}{\partial x^i} \right\|_{L^2(\lambda_n)} \frac{1}{2^n}$$

where $\lambda_n(x) = \gamma_n(x)dx$, with γ_n the gaussian density with covariance matrix $\delta_{ij}2^{-n}, i, j = 0, ..., 2^n$. With this definition one has $\|F\|_{1,2} = \|f_n\|_{n,1,2}$ if $F \in S_n$ and it's representation function is f_n. So the closure of S_n with respect to $\|\circ\|_{1,2}$ corresponds to the closure of C_n^∞ with respect to $\|\circ\|_{n,1,2}$ - which replaces the norm $\|\circ\|_{D,1,2}$ from the standard Sobolev spaces. This is the correspondence at the level n. But Malliavin's calculus is a calculus in infinite dimension and this extension is obtained by means of the embedding in (16) and (19). The central point here is that for $f \in C_n^\infty$ one has $\|j_n(f)\|_{n+1,1,2} = \|f \circ s_n\|_{n+1,1,2} = \|f\|_{n,1,2}$ (this is because $\Delta_n^k \sim \Delta_{n+1}^{2k} + \Delta_n^{2k+1}$). So one may define a norm $\|\circ\|_{\infty,1,2}$ on the whole $C_\infty^\infty = \cup_{n \in N} C_n^\infty$ which coincides with $\|\circ\|_{n,1,2}$ on each C_n^∞. Then the space $D^{1,2}$ corresponds to the closure of C_∞^∞ with respect to $\|\circ\|_{\infty,1,2}$. Of course this space is much larger then the closure of C_n^∞ with respect to $\|\circ\|_{n,1,2}$. This is true even if we consider C_n^∞ as a subspace of C_∞^∞, because the completation is done using sequences which are not contained in a single sub-space C_n^∞. And this is the reason for which the Malliavin calculus is a really infinite dimensional differential calculus and not just a collection of finite dimensional ones.

We introduce now higher order derivatives in a similar way. For a multi-index $\alpha = (\alpha_1, ..., \alpha_r) \in \{1, ..., d\}^r$ we denote $|\alpha| = r$ and we define

$$D^\alpha_{(t_1,\dots,t_r)} F = \frac{\partial^r F}{\partial \Delta_{t_1}^{\alpha_1} ... \partial \Delta_{t_r}^{\alpha_r}} = \sum_{k_1,\dots,k_r=0}^{m-1} \frac{\partial^r f}{\partial x_n^{k_1,\alpha_1} ... \partial x_n^{k_r,\alpha_r}} (\Delta_n^0, ..., \Delta_n^{m-1}) \prod_{j=1}^r 1_{I_n^{k_j}}(t_j) \quad (25)$$

So D^α is an operator from S to the space of simple processes with r time-parameters, which is a dense linear subspace of $L^2([0,\infty)^r \times \Omega, B_+^{\otimes r} \times F, \lambda^{\otimes r} \times P)$. One proves in a similar way that D^α is closable and defines its extension to $Dom(D^\alpha)$.

We introduce the following norms on S :

$$\|F\|_{r,p}^p = E\,|F|^p + \sum_{q=1}^r \sum_{|\alpha|=q} E \int_0^\infty ... \int_0^\infty \left| D^\alpha_{(t_1,\dots,t_q)} F \right|^p dt_1 ... dt_q \quad (26)$$

and we define $D^{r,p}$ to be the closure of S with respect to $\|\circ\|_{r,p}$. We define

$$D^\infty = \cap_{r=1}^\infty \cap_{p \geq 1} D^{r,p}. \quad (27)$$

Finally we define the so called Orenstein Ulemback operator $L : S \to S$ by

$$L(F) = -\sum_{i=1}^d \delta^i (D^i F). \quad (28)$$

If $F = f(\Delta_n^0, ..., \Delta_n^{m-1})$ then we have

$$L(F) = \sum_{i=1}^{d}\sum_{k=0}^{m-1} \frac{\partial^2 f}{\partial (x^{k,i})^2}(\Delta_n)\frac{1}{2^n} - \sum_{i=1}^{d}\sum_{k=0}^{m-1} \frac{\partial f}{\partial x^{k,i}}(\Delta_n)\Delta_n^{k,i}. \tag{29}$$

As an immediate consequence of the duality between D^i and δ^i one has $E(FL(G)) = E(GL(F))$ for every $F, G \in S$. This permits to prove that $L : S \subset L^2(\Omega, F_\infty, P) \to S \subset L^2(\Omega, F_\infty, P)$ is closable and to define its extension to $Dom(L) \subset L^2(\Omega, F_\infty, P)$. Using approximation with simple functionals one may also obtain

$$E(FL(G)) = E(GL(F)) = -E(< DF, DG > \tag{30}$$

for $F, G \in Dom(L) \cap D^{1,2}$ (in fact $D^{1,2} \subset Dom(L)$ but this is not still clear at this stage - see the following section). We give now some simple properties of the above operators.

Proposition 7 *i) Let $\phi \in C_p^1(R^k; R)$ and $F = (F^1, ..., F^k)$ with $F^i \in D^{1,2}$. Then $\phi(F) \in D^{1,2}$ and for each $i = 1, ..., d$*

$$D_t^i \phi(F) = \sum_{j=1}^{k} \frac{\partial \phi}{\partial x^j}(F) D^i F^j. \tag{31}$$

ii) Let $F, G \in Dom(L) \cap D^{1,4}$ be such that $FG \in Dom(L)$. Then

$$L(FG) = FLG + GLF + 2 < DF, DG > . \tag{32}$$

More generally, if $F = (F^1, ..., F^d)$ with $F^i \in Dom(L) \cap D^{1,4}, i = 1..., d$ and $\phi \in C_p^2(R^d; R)$, then $\phi(F) \in Dom(L)$ and

$$L\phi(F) = \sum_{i,j=1}^{d} \frac{\partial^2 \phi}{\partial x^i \partial x^j}(F) < DF^i, DF^j > + \sum_{i=1}^{d} \frac{\partial \phi}{\partial x^i}(F) LF^i.$$

The proof of these formulas follow from the standard properties of differential calculus in the case of simple functionals and then extends to general functionals by approximation with simple functionals.

We denote $L_a^{1,2}$ the space of the adapted square integrable processes $u = u(t, \omega)$ such that for each fixed $t, u(t, \omega) \in D^{1,2}$ and there exists a measurable version of the double indexed process $D_s u(t, \omega)$ which is square integrable, that is $E\int_0^\infty \int_0^\infty \sum_{i=1}^d \left|D_s^i u(t, \omega)\right|^2 dsdt < \infty$.

Proposition 8 *Let $u \in L_a^{1,2}$. Then*

$$\begin{aligned} D_t^i(\int_0^\infty u(s,\omega) dB_s^j) &= u(t,\omega)\delta_{ij} + \int_t^\infty D_t^i u(s,\omega) dB_s^j \\ D_t^i(\int_0^\infty u(s,\omega) ds) &= \int_t^\infty D_t^i u(s,\omega) ds \end{aligned} \tag{33}$$

with δ_{ij} the Kronecker symbol.

The proof is easy for simple functionals and extends by approximation for general functionals.

We close this section with the Clark-Ocone formula which gives an explicit expression of the density in the martingale representation theorem in terms of Malliavin derivatives. This is especially interesting in mathematical finance where this density represents the strategy.

Theorem 1 *Let $F \in D^{1,2}$. Then*

$$F = EF + \sum_{i=1}^{d} \int_0^\infty E(D_s^i F \mid F_s) dB_s^i \quad a.s \tag{34}$$

Proof. We assume that $EF = 0$ (if not we take $F - EF$). We fix a constant c and wr take d simple processes $u^i, i = 1, ..., d$ which are previsible and we employ the integration by parts formula and the isometry property for stochastic integrals and obtain:

$$\begin{aligned} EF \times \left(c + \sum_{i=1}^{d} \int_0^\infty u_s^i dB_s^i \right) &= E(F \sum_{i=1}^{d} \delta^i(u^i)) \\ &= \sum_{i=1}^{d} E \int_0^\infty D_s^i F \times u_s^i ds = \sum_{i=1}^{d} E \int_0^\infty E(D_s^i F \mid F_s) \times u_s^i ds \\ &= \sum_{i=1}^{d} E \int_0^\infty E(D_s^i F \mid F_s) dB_s^i \times \int_0^\infty u_s^i dB_s^i \\ &= E \left(\sum_{i=1}^{d} \int_0^\infty E(D_s^i F \mid F_s) dB_s^i \right) \times \left(c + \sum_{i=1}^{d} \int_0^\infty u_s^i dB_s^i \right). \end{aligned}$$

The representation theorem for martingales guarantees that the random variables of the form $c + \sum_{i=1}^{d} \int_0^\infty u_s^i dB_s^i$ are dense in $L^2(\Omega)$ and so we obtain $F = \sum_{i=1}^{d} \int_0^\infty E(D_s^i F \mid F_s) dB_s^i, a.s.$ $\square$

3.2 The integration by parts formula and the absolute continuity criterion

The aim of this section is to employ the differential operators constructed in the previous and the duality between D and δ in order to obtain an integration by parts formula of type (1) and establish the corresponding absolute continuity criterion. Given a random vector $F = (F^1, ..., F^n)$ with $F^i \in D^{1,2}, i = 1, ..., n$, one defines Malliavin's covariance matrix $\gamma_F = (\gamma_F^{ij})_{i,j=1,n}$ by

$$\gamma_F^{ij} =< DF^i, DF^j >= \sum_{r=1}^{d} \int_0^\infty D_t^r F^i D_t^r F^j dt, \quad i,j = 1,...,n. \tag{35}$$

In the gaussian case Malliavin's covariance matrix coincides with the classical covariance matrix of the gaussian variable. In fact, let $\sigma_j^i, i = 1,...,n, j = 1,...,d$ be a deterministic matrix and let $X_t^i = \sum_{j=1}^d \sigma_j^i B_t^j, i = 1,...,n$. Then $F = X_t$ is an $n-$dimensional centered gaussian random variable with covariance matrix $t \times \sigma\sigma^*$. On the other hand this is a simple functional and $D_s^r X_t^i = 1_{[0,t]}(s)\sigma_r^i$ so that $\sum_{r=1}^d \int_0^\infty D_s^r X_t^i D_s^r X_t^j ds = t \times \sum_{r=1}^d \sigma_r^i \sigma_r^j = t \times (\sigma\sigma^*)_j^i$. Recall now that we want to obtain an absolute continuity criterion (with respect to the Lebesque's measure) for the law of F. But if $t \times \sigma\sigma^*$ is degenerated then $F = X_t$ lives (in the sense of the support of the law) in some subspace of R^n of dimension $m < n$ and so may not be absolutely continuous with respect to the Lebesgue's measure on R^n. In fact the law of a gaussian random variable is absolutely continuous with respect to the Lebesgue's measure if and only if the covariance matrix is not degenerated. So it appears as natural to consider the following non-degeneracy property:

$$(H_F) \quad E((\det \gamma_F)^{-p}) < \infty, \quad \forall p > 1. \tag{36}$$

This property says that $\det \gamma_F \neq 0$ *a.s.*, but it says much more, because $(\det \gamma_F)^{-p}$ has to be integrable for every $p > 0$. In fact Boulau and Hirsch [5] proved that if $F \in (D^{1,2})^n$ and $\det \gamma_F \neq 0$ *a.s.* then the law of F is absolutely continuous with respect to the Lebesgue's measure, but one has no regularity for the density. The stronger assumption (H_F) will permit to obtain a smooth density and to give evaluations of this density and of its derivatives as well.

Let us go on and give the integration by parts formula.

Theorem 2 *Let $F = (F^1, ..., F^n)$ with $F^i \in D^{2,2}, i = 1,...,n$. Suppose that (H_F) holds true. Then, for every $G \in D^{1,2}$ and every $\phi \in C_p^1(R^n; R)$*

$$E(\frac{\partial \phi}{\partial x^i}(F)G) = E(\phi(F)H^{(i)}(F;G)) \tag{37}$$

with

$$H^{(i)}(F;G) = -\sum_{j=1}^{n} (G\widehat{\gamma}_F^{ji} L(F^j) + < DF^j, D(G\widehat{\gamma}_F^{ji}) >) \tag{38}$$

with $\widehat{\gamma}_F = \gamma_F^{-1}$.

Moreover, if $F^i \in D^\infty, i = 1,...,n$ and $G \in D^\infty$ then for every multi-index $\alpha = (\alpha_1, ..., \alpha_m) \in \{1, ..., d\}^m$

$$E(D^\alpha \phi(F)G) = E(\phi(F)H^\alpha(F : G)) \tag{39}$$

with H^α defined by the recurrence formulas

$$H^{(i,\alpha)}(F;G) = H^{(i)}(F; H^\alpha(F;G)) \tag{40}$$

where $(i,\alpha)=(i,\alpha_1,...,\alpha_m)$ *for* $\alpha=(\alpha_1,...,\alpha_m)$.

In the ennonce of the above proposition we apply the operator L to F^i. So we have to know that $F^i\in Dom(L)$. We will prove in the next section that $Dom(L)=D^{2,2}$ and so the above opertation is legitime.

Proof. Using the chain rule (31)

$$\begin{aligned} < & D\phi(F),DF^j>=\sum_{r=1}^{d}<D^r\phi(F),D^rF^j> \\ = & \sum_{r=1}^{d}\sum_{i=1}^{n}\frac{\partial\phi}{\partial x^i}(F)<D^rF^i,D^rF^j>=\sum_{i=1}^{n}\frac{\partial\phi}{\partial x^i}(F)\gamma_F^{ij}\end{aligned}$$

so that

$$\frac{\partial\phi}{\partial x^i}(F)=\sum_{j=1}^{n}<D\phi(F),DF^j>\widehat{\gamma}_F^{ji}.$$

Moreover, using (32) $<D\phi(F),DF^j>=\frac{1}{2}(L(\phi(F)F^j)-\phi(F)L(F^j)-F^jL(\phi(F)))$. Then using the duality relation (30)

$$\begin{aligned} E(\frac{\partial\phi}{\partial x^i}(F)G) &= \frac{1}{2}\sum_{j=1}^{n}E(G(L(\phi(F)F^j)-\phi(F)L(F^j)-F^jL(\phi(F))\widehat{\gamma}_F^{ji}) \\ &= \frac{1}{2}\sum_{j=1}^{n}E(\phi(F)(F^jL(G\widehat{\gamma}_F^{ji})-G\widehat{\gamma}_F^{ji}L(F^j)-L(GF^j\widehat{\gamma}_F^{ji})) \\ &= \frac{1}{2}\sum_{j=1}^{n}E(\phi(F)(F^jL(G\widehat{\gamma}_F^{ji})-G\widehat{\gamma}_F^{ji}L(F^j)-L(GF^j\widehat{\gamma}_F^{ji})) \\ &= -\sum_{j=1}^{n}E(\phi(F)(G\widehat{\gamma}_F^{ji}L(F^j)+<DF^j,D(G\widehat{\gamma}_F^{ji})>).\end{aligned}$$

The relation (39) is obtained by recurrence.. $\square$

As an immediate consequence of the above integration by parts formula and of Proposition 4 we obtain the following result.

Theorem 3 *Let* $F=(F^1,...,F^n)$ *with* $F^i\in D^\infty, i=1,...,n$. *Suppose that* (H_F) *holds true.* $P\circ F^{-1}(dx)=p_F(x)dx$ *Moreover* $p_F\in C^\infty(R^n;R)$ *and for each multi-index* α *and each* $k\in N, \sup_{x\in R^n}|x|^k|D^\alpha p_F(x)|<\infty$.

3.2.1 On the covariance matrix

Checking the non - degeneracy hypothesis (H_F) represents the more difficult problem when using the integration by parts formula in concrete situations. We give here a general lemma which reduces this problem to a simpler one.

Lemma 4 *Let γ be a random $n \times n$ dimensional matrix which is non negative defined and such that $(E\left|\gamma^{ij}\right|^p)^{1/p} \leq C_p < \infty, i,j = 1,...,n$. Suppose that for each $p \in N$ there exists some $\varepsilon_p > 0$ and $K_p < \infty$ such that for every $0 < \varepsilon < \varepsilon_p$ one has*

$$\sup_{|\xi|=1} P(<\gamma\xi,\xi><\varepsilon) \leq K_p \varepsilon^p. \tag{41}$$

Then $(\det \gamma)^{-1}$ has finite moments of any order and more precisely

$$E\left|\det \gamma\right|^{-p} \leq 2\delta_{n,np+2} \tag{42}$$

where $\delta_{n,p} =: (c_n K_{p+2n} + 2^p(n+1)^p C_p)$ with c_n an universal constant which depends on the dimension n.

Proof. Step1. We shall first prove that for every $p > 2n$

$$P(\inf_{|\xi|=1} <\gamma\xi,\xi><\varepsilon) \leq (c_n K_{p+2n} + 2^p(n+1)^p C_p)\varepsilon^p. \tag{43}$$

We take $\xi_1,...,\xi_N \in S_{n-1} =: \{\xi \in R^n : |\xi| = 1\}$ such that the balls of centers ξ_i and radius ε^2 cover S_{n-1}. One needs $N \leq c_n\varepsilon^{-2n}$ points, where c_n is a constant depending on the dimension. It is easy to check that $|<\gamma\xi,\xi> - <\gamma\xi_i,\xi_i>| \leq (n+1)\left|\gamma\right|\left|\xi-\xi_i\right|$ where $\left|\gamma\right|^2 = \sum_{i,j=1}^d \left|\gamma^{ij}\right|^2$. It follows that $\inf_{|\xi|=1} <\gamma\xi,\xi> \geq \inf_{i=1,N} <\gamma\xi_i,\xi_i> -(n+1)\left|\gamma\right|\varepsilon^2$ and consequently

$$\begin{aligned}
P(\inf_{|\xi|=1} &< \gamma\xi,\xi><\varepsilon) \leq P(\inf_{i=1,N} <\gamma\xi_i,\xi_i><\frac{\varepsilon}{2}) + P(\left|\gamma\right| \geq \frac{1}{2(n+1)\varepsilon}) \\
&\leq N \max_{i=1,N} P(<\gamma\xi_i,\xi_i><\frac{\varepsilon}{2}) + 2^p(n+1)^p\varepsilon^p E\left|\gamma\right|^p \\
&\leq c_n K_p \varepsilon^{p-2n} + 2^p(n+1)^p\varepsilon^p C_p
\end{aligned}$$

and so (43) is proved.

Step 2. $\lambda = \inf_{|\xi|=1} <\gamma\xi,\xi>$ is the smaller proper value of the matrix γ so that $\det\gamma \geq \lambda^n$ and consequently $(E\left|\det\gamma\right|^{-p})^{1/p} \leq (E\lambda^{-np})^{1/p}$. On the other hand we know from (43) that $P(\lambda^{-1} \geq \frac{1}{\varepsilon}) \leq \delta_{n,p}\varepsilon^p$ with $\delta_{n,p} =: (c_n K_{p+2n} + 2^p(n+1)^p C_p)$. Taking $\varepsilon_k = \frac{1}{k}$ we obtain

$$\begin{aligned} E\lambda^{-np} &= \sum_{k=0}^{\infty} E(\lambda^{-np} 1_{\{\lambda^{-1}\in[k,k+1)\}}) \leq \sum_{k=0}^{\infty} (\bar{k}+1)^{np} P(\lambda^{-1} \geq k) \\ &\leq \sum_{k=0}^{\infty} (k+1)^{np} \delta_{n,np+2} \frac{1}{k^{np+2}} \leq 2\delta_{n,np+2} \end{aligned}$$

and the proof is complete. □

We give a first application of this lemma in a frame which roughly speaking corresponds to the ellipticity assumption for diffusion processes. Let $F = (F^1, ..., F^n)$ with $F^i \in D^{1,2}$. Assume that

$$D_s^l F^i = q_l^i(s,\omega) + r_l^i(s,\omega)$$

where q_l^i and r_l^i are measurable stochastic processes which satisfy the following assumptions.

There exists a positive random variable a such that

$$\begin{aligned} i) \quad \sum_{l=1}^{d} (\sum_{i=1}^{d} q_l^i(s,\omega)\xi^i)^2 &\geq a\,|\xi|^2 \quad ds \times dP(\omega) \ a.s. \\ ii) \quad P(a &\leq \varepsilon) \leq K_p \varepsilon^p, \quad \forall p \in N, \varepsilon > 0 \end{aligned} \tag{44}$$

for some constants K_p.

There exists some $\delta > 0$ and some constants C_p such that

$$E \sup_{s\leq\varepsilon} \left| r_l^{ij}(s) \right|^p \leq C_p \varepsilon^{p\delta} \quad \forall p \in N, \varepsilon > 0. \tag{45}$$

Remark 6 *Let $Q^{ij} = \sum_{l=1}^{d} q_l^i q_l^j$. Then $\sum_{l=1}^{d} (\sum_{i=1}^{d} q_l^i(s,\omega)\xi^i)^2 =< Q\xi, \xi >$ so i) represents the "ellipticity assumption" for Q, at list if a is a strictly positive constant. If a is not a constant but a random positive variable then $a > 0$ has to be replaced by $Ea^{-p} < \infty, \forall p \in N$, and this is equivalent with ii). On the other hand (45) says that r_s may be ignored, at list for small s.*

Lemma 5 *If (44) and (45) hold true then $E\,|\det \gamma_F|^{-p} < \infty, \forall p \in N$, and so (H_F) holds true.*

Proof. We will employ the previous lemma so we check (41).

Le $\beta = 1 - \delta$ and let $\xi \in R^n$ be an unitary vector. We write

$$\begin{aligned} < \quad \gamma_F \xi, \xi >= \sum_{l=1}^{d} \int_0^t < D_s^l F, \xi >^2 ds \geq \sum_{l=1}^{d} \int_0^{\varepsilon^\beta} < D_s^l F, \xi >^2 ds \\ \geq \quad \int_0^{\varepsilon^\beta} (a - \sum_{l=1}^{d} |r_l(s)|^2) ds \geq (a - \sum_{l=1}^{d} \sup_{s\leq\varepsilon^\beta} |r_l(s)|^2) \varepsilon^\beta. \end{aligned}$$

It follows that

$$\begin{aligned}
P(&< \gamma_F \xi, \xi >< \varepsilon) \leq P(a - \sum_{l=1}^{d} \sup_{s \leq \varepsilon^\beta} |r_l(s)|^2 < \varepsilon^{1-\beta}) \\
&\leq P(a < 2\varepsilon^{1-\beta}) + P(\sum_{l=1}^{d} \sup_{s \leq \varepsilon^\beta} |r_l(s)|^2 \geq \varepsilon^{1-\beta}) \\
&\leq K_p (2\varepsilon^{1-\beta})^p + \frac{1}{\varepsilon^{p(1-\beta)}} C_{2p} \varepsilon^{2p\delta} = (K_p 2^p + C_p)\varepsilon^{\delta p}
\end{aligned}$$

and the proof is completed. □

The idea in the above lemma is clear: one decomposes DF in a principal term q which is non-degenerated and a rest $r(s)$ which is small as $s \searrow 0$. Then the main trick is to localize around zero (one replaces $\int_0^t$ by $\int_0^{\varepsilon^\beta}$) in order to distinguish between q and r. The same idea works in a more sophisticated frame when the non degeneracy assumption is weaker then ellipticity. But it seems difficult to describe such a non degeneracy condition in an abstract frame - in the following section such a condition is given in the particular frame of diffusion processes. What is specific in this frame is that if $F = X_t$ then one has the decomposition $D_s F = D_s X_t = Y_t \rho_s$ so that $\gamma_{X_t} = Y_t \int_0^t \rho_s \rho_s^* ds Y_t^*$. Moreover Y_t is an invertible matrix and so the invertibility of γ_{X_t} is equivalent with the invertibility of $\overline{\gamma}_t = \int_0^t \rho_s \rho_s^* ds$. Although the above decomposition is specific to diffusion processes and it is not clear that there are other interesting examples, we will now discuss the invertibility of matrixes of the form $\overline{\gamma}_t$. The main tool in this analysis is the decomposition of ρ in stochastic series (such ideas have been developed in [2], [12],[1]).

Let us introduce some notation. For a multi-index $\alpha = (\alpha_1, ..., \alpha_m) \in \{0, 1, ..., d\}^m$ we denote $p(\alpha) = \frac{1}{2} card\{i : \alpha_i \neq 0\} + card\{\alpha_i = 0\}$. In order to get unitary notation we denote $B_t^0 = t$ so that $dB_t^0 = dt$ and, for an previsible square integrable process U we define the multiple integral $I_\alpha(U)(t) = \int_0^t \int_0^{t_1} ... \int_0^{t_{m-1}} U(t_m) dB_{t_m}^{\alpha_m} ... dB_{t_1}^{\alpha_1}$. Note that this integral is different from the ones which appear in the chaos decomposition for several reasons. First of all it does not contain just stochastic integrals but Lebesgues integrals as well. Moreover, U is not a deterministic function of m variables but a stochastic process. So the decomposition in stochastic series is compleatly different from the one in Wiener chaos.

We say that a stochastic process ρ admits a decomposition in stochastic series of order m if there exists some constants c_α and some previsible square integrable processes U_α such that

$$\rho_t = \sum_{p(\alpha) \leq m/2} c_\alpha I_\alpha(1)(t)) + \sum_{p(\alpha) = (m+1)/2} I_\alpha(U_\alpha)(t).$$

We employ the convention that for the void multi-index we have $p(\emptyset) = 0$ and $I_0(1)(t) = 1$. The central fact concerning stochastic seres is the following.

Lemma 6 *i) Let $m \in N$ and let $c = (c_\alpha)_{p(\alpha)\leq m/2}$. We denote $|c|^2 = \sum_{p(\alpha)\leq m/2} c_\alpha^2$. Then, for every $\beta < \frac{1}{m+1}$ and every $p \in N$ there exists a constant $C_{p,\beta}$ such that*

$$\inf_{|c|\geq 1} P(\int_0^{\varepsilon^\beta} (\sum_{p(\alpha)\leq m/2} c_\alpha I_\alpha(1)(s))^2 ds < \varepsilon) \leq C_{\beta,p}\varepsilon^p. \tag{47}$$

ii) Let $U = (U_\alpha)_{p(\alpha)=(m+1)/2}$ be such that $E\sup_{s\leq t} |U_\alpha|^p < \infty, \forall t > 0, p \in N$. Then

$$(E \sup_{s\leq t} \left| \sum_{p(\alpha)=(m+1)/2} I_\alpha(U_\alpha)(s)) \right|^{-p})^{1/p} \leq K_p t^{(m+1)/2} < \infty. \tag{48}$$

Roughly speaking i) says that if c is not degenerated then $\sum_{p(\alpha)\leq m/2} c_\alpha I_\alpha(1)(s)$ is not degenerated. This property is analogues with (44). The proof is non trivial and we do not give it here (see [12] Theorem $(A.6)$ $pg61$, or [1]). ii) says that $\sum_{p(\alpha)=k/2} I_\alpha(U_\alpha)(s))$ vanishes as $t \to 0$ as $t^{(m+1)/2}$ and this will be used in order to control the remainder of the series. (48) is an easy application of Burckholder and Holder inequalities so we it live out.

We consider now a $n \times d$ dimensional matrix $\rho(s) = (\rho^{ij}(s))_{i,=1,n,j=1,d}$ such that

$$\rho^{ij}(s) = \sum_{k=0}^{m} c_\alpha^{ij} I_\alpha(1)(s) + \sum_{p(\alpha)=(m+1)/2} I_\alpha(U_\alpha^{ij})(s)$$

and define

$$Q_m(c) = \inf_{|\xi|=1} \sum_{p(\alpha)\leq m/2} < c_\alpha c_\alpha^* \xi, \xi >= \inf_{|\xi|=1} \sum_{l=1}^{d} \sum_{p(\alpha)\leq m/2} < c_\alpha^l, \xi >^2 .$$

We are interested in the $n \times n$ dimensional matrix $\gamma_t = (\gamma_t^{ij})_{i,j=1,n}$ defined by $\gamma_t^{ij} = \int_0^t (\rho_s \rho_s^*)^{ij} ds.$

Proposition 9 *If $Q_m(c) > 0$ then for every $t \geq 0$ the matrix γ_t is invertible and $E(\det \gamma_t)^{-p} < \infty$.*

Proof. We will use lemma 19 so we have to check the hypothesis (41). We denote $q_s^{ij} = \sum_{p(\alpha)\leq m} c_\alpha^{ij} I_\alpha(1)(s)$ and $r_s^{ij} = \sum_{p(\alpha)=(m+1)/2} I_\alpha(U_\alpha^{ij})(s)$ so that $\gamma_t = \int_0^t (q_s + r_s) ds$.

We employ now the same strategy as in the proof of lemma 20. We take $\beta > 0$ to be chosen latter on. For a unitary vector ξ we have

$$\begin{aligned}
< \quad \gamma_t \xi, \xi >&= \int_0^t \sum_{l=1}^d < q_s^l + r_s^l, \xi >^2 ds \geq \int_0^{\varepsilon^\beta} \sum_{l=1}^d < q_s^l + r_s^l, \xi >^2 ds \\
\geq & \quad \frac{1}{2} \int_0^{\varepsilon^\beta} \sum_{l=1}^d < q_s^l, \xi >^2 ds - \int_0^{\varepsilon^\beta} \sum_{l=1}^d < r_s^l, \xi >^2 ds \\
\geq & \quad \frac{1}{2} \int_0^{\varepsilon^\beta} \sum_{l=1}^d < q_s^l, \xi >^2 ds - \varepsilon^\beta \sup_{s \leq \varepsilon^\beta} |r_s|^2
\end{aligned}$$

with $|r_s|^2 = \sum_{i,j} \left| r_s^{ij} \right|^2$.
Then

$$P(< \gamma_t \xi, \xi >< \varepsilon) \leq P(\int_0^{\varepsilon^\beta} \sum_{l=1}^d < q_s^l, \xi >^2 ds < 4\varepsilon) + P(\sup_{s \leq \varepsilon^\beta} |r_s|^2 > \varepsilon^{1-\beta}).$$

Using (48) the second term is dominated by $C_p \varepsilon^{p((m+2)\beta-1)}, \forall p \in N$. It follows that, if $\beta \geq \frac{1}{m+2}$ then $\lim_{\varepsilon \to 0} \varepsilon^{-k} P(\sup_{s \leq \varepsilon^\beta} |r_s|^2 > \varepsilon^{1-\beta}) = 0$ for every $k \in N$.
The first term is dominated by

$$\min_{l=1,d} P(\int_0^{\varepsilon^\beta} < q_s^l, \xi >^2 ds < 4\varepsilon).$$

Note that $\int_0^{\varepsilon^\beta} < q_s^l, \xi >^2 ds = \int_0^{\varepsilon^\beta} (\sum_{p(\alpha) \leq m} \sum_{i=1}^n c_\alpha^{il} \xi^i I_\alpha(1)(s))^2 ds$. At list for one $l = 1, ..., d$ one has $\sum_{p(\alpha) \leq m} \sum_{i=1}^n \left| c_\alpha^{il} \xi^i \right|^2 \geq \frac{1}{d} Q_m(c) > 0$. Then, if $\beta < \frac{1}{m+1}$, (47) yields $\lim_{\varepsilon \to 0} \varepsilon^{-k} \inf_{|\xi|=1} P(\int_0^{\varepsilon^\beta} < q_s^l, \xi >^2 ds < 4\varepsilon) = 0$ for each $k \in N$. So we take $\frac{1}{m+2} < \beta < \frac{1}{m+1}$ and the proof is completed. □

3.3 Wiener chaos decomposition

Malliavin calculus is intimately related to the Wiener chaos decomposition. On one hand one may define the differential operators using directly this decomposition (this is the approach in [O]) and give a precise description of their domains. On the other hand this is the starting point of the analysis on the Wiener space which has been developed in [15],[20],[10]... and so on. But, as it is clear from the previous section, one may also set up the Malliavin calculus and use this calculus in concrete applications without even mentioning the chaos decomposition. This is why we restric ouerself to a short insight to this topoic: we just give the charactherisation of the domain of the differential operators and a relative compactness criterios on the Wiener space, but live out more involved topics as Meyer's inequalities and the distribution theory on the Weiner space.

In order to simplifying the notation we just consider an one dimensional Brownian motion and we work on the time interval $[0,1]$ (in particular the simple functionals are of the form $F=f(\Delta_n^0,...,\Delta_n^{2^n-1}))$. The ideas are quit the same in the general frame. We denote $L^2_{s,k}$ the subspace of the symmetric functions which belong to $L^2([0,1]^k)$. For $f_k\in L^2_{s,k}$ we define the itterated integral

$$\int_0^1\int_0^{t_1}...\int_0^{t_{k-1}} f_k(t_1,...,t_k)dB_{t_k}...dB_{t_k}$$

and the so called multiple stochastic integral

$$I_k(f_k)=k!\int_0^1\int_0^{t_1}...\int_0^{t_{k-1}} f_k(t_1,...,t_k)dB_{t_k}...dB_{t_k}.$$

$I_k(f_k)$ may be considered as the integral of f_k on the whole cube $[0,1]^k$ while the iterated integral is just the integral on a simplex.

As an immediate consequence of the isometry property and of the fact that the expectation of stochastic integrals is zero one obtains

$$\begin{aligned} E(I_k(f_k)I_p(f_p)) &= k!\,\|f_k\|^2_{L^2[0,1]^k} \quad if \quad k=p \\ &= 0 \quad if \quad k\neq p. \end{aligned}$$

One defines the chaos of order k to be the space

$$H_k=\{I_k(f_k):f_k\in L^2_{s,k}\}.$$

It is easy to check that H_k is a closed linear subspace of $L^2(\Omega)$ and H_k is orthogonal to H_p for $k\neq p$. For $k=0$ one defines $H_0=R$, so the elements of H_0 are constants. We have the Wiener chaos decomposition

$$L^2(\Omega)=\bigoplus_{k=0}^{\infty}H_k.$$

We give not the proof here - see [15] for example. The explicit expression of the above decomposition is given in the following theorem.

Theorem 4 *For every $F\in L^2(\Omega)$ there exists a unique sequence $f_k\in L^2_{s,k},k\in N$ such that*

$$\begin{aligned} \sum_{k=0}^{\infty} k!\,\|f_k\|^2_{L^2[0,1]^k} &= \|F\|^2_{L^2(\Omega)}<\infty \quad and \\ F &= \sum_{k=0}^{\infty} I_k(f_k). \end{aligned} \tag{49}$$

Let us now see which is the relation between multiple integrals and simple functionals. Let $n \in N$ be fixed. For a multi-index $\overline{i} = \{i_1, ..., i_k\} \in \Theta_k^n =: \{0, 1, ..., 2^n - 1\}^k$ we denote $\Gamma_{\overline{i}}^n = [t_n^{i_1}, t_n^{i_1+1}) \times .... \times [t_n^{i_k}, t_n^{i_k+1})$ and for $x = (x^0, ..., x^{2^n-1})$ we define $x^{\overline{i}} = \prod_{j=1}^k x^{i_j}$. Let $q = (q_{\overline{i}})_{\overline{i} \in \Theta_k^n}$ be a $k-$dimensional matrix which is symmetric ($q_{\overline{i}} = q_{\pi(\overline{i})}$ for every permutation π) and which has null elements on the diagonals, that is $q_{\overline{i}} = 0$ if $i_k = t_p$ for some $k \neq p$. We denote by Q_k^n the set of all the matrixes having these properties. To a matrix $q \in Q_k^n$ we associate the piecewise constant, symmetric, function

$$f_{k,q}(t_1, ..., t_k) = \sum_{\overline{i} \in \Theta_k^n} q_{\overline{i}} 1_{\Gamma_{\overline{i}}^n}(t_1, ..., t_k)$$

so that $f_{k,q} = q_{\overline{i}}$ on $\Gamma_{\overline{i}}^n$. We also define the polynomial

$$P_q(x^0, ..., x^{2^n-1}) = \sum_{\overline{i} \in \Theta_k^n} q_{\overline{i}} x^{\overline{i}}.$$

Note that $(\Delta_n)^{\overline{i}} = \prod_{j=1}^k \Delta_n^{i_j} = \prod_{j=1}^k (B(t_n^{i_j+1}) - B(t_n^{i_j}))$ appears as a "gaussian measure" of the cube $\Gamma_{\overline{i}}^n$. Of course one can not extend such a measure to all the Borel sets of $[0,1]^k$ because the Brownian path has not finite variation, and this is why stochastic integrals come in. Then $P_{k,q}(\Delta_n)$ appears as the integral of $f_{k,q}$ with respect to this measure. Moreover we have

$$I_k(f_{k,q}) = P_{k,q}(\Delta_n). \tag{50}$$

Let us check the above equality in the case $k = 2$ (the proof is the same in the general case). We write

$$\begin{aligned} I_k(f_{k,q}) &= 2! \int_0^1 \int_0^{s_1} f_{k,q}(s_1, t_{s_2}) dB_{s_2} dB_{s_1} = 2 \sum_{i=0}^{2^n-1} \int_{t_n^i}^{t_n^{i+1}} \int_0^{s_1} f_{k,n}(t_n^i, s_2) dB_{s_2} dB_{s_1} \\ &= 2 \sum_{i=0}^{2^n-1} \int_{t_n^i}^{t_n^{i+1}} \sum_{j=0}^{i-1} \int_{t_n^j}^{t_n^{j+1}} f_{k,n}(t_n^i, t_n^j) dB_{s_2} dB_{s_1} + 2 \sum_{i=0}^{2^n-1} \int_{t_n^i}^{t_n^{i+1}} \int_{t_n^i}^{s_1} f_{k,n}(t_n^i, t_n^i) dB_{s_2} dB_{s_1}. \end{aligned}$$

Since $f_{k,n}$ is null on the diagonal the second term is null. So we obtain

$$I_k(f_{k,q}) = 2 \sum_{i=0}^{2^n-1} \sum_{j=0}^{i-1} f_{k,n}(t_n^i, t_n^j) \int_{t_n^i}^{t_n^{i+1}} \int_{t_n^j}^{t_n^{j+1}} 1 dB_{s_2} dB_{s_1} = P_{k,q}(\Delta_n)$$

the last equality being a consequence of the symmetry.. So (50) is proved. As a by product we obtain the following result:

Lemma 7 *i) For every $f_k \in L^2_{s,k}$, $I_k(f_k) \in D^\infty \cap Dom(L), (I_{k-1}(f_k(s,\circ)))_{s\geq 0} \in Dom(\delta)$, and*

$$\begin{aligned} a) \quad D_s I_k(f_k) &= I_{k-1}(f_k(s,\circ)), \quad \forall k \geq 1, \\ b) \quad D^{(p)}_{s_1,\ldots,s_p} I_k(f_k) &= I_{k-p}(f_k(s_1,\ldots,s_p,\circ)), \quad \forall k \geq p, \\ c) \quad \delta(I_{k-1}(f_k(s,\circ))) &= I_k(f_k), \\ d) \quad L I_k(f_k) &= -k I_k(f_k). \end{aligned} \tag{51}$$

ii) $Sp\{P_{k,q}(\Delta_n) : q \in Q^n_k, k,n \in N\}$ is dense in $L^2(\Omega)$. In particular one may take $P_{k,q}(\Delta_n) : q \in Q^n_k, k,n \in N$ as the initial class of simple functionals.

Proof. Let us prove ii) first. The symmetric piecewise constant functions $f_{k,q}, q \in \Theta^n_k, n \in N$, are dense in $L^2_{s,k}$ so $P_{k,q}(\Delta_n), q \in \Theta^n_k, n \in N$ are dense in H_k (one employes the isometry property in order to check it). Now ii) follows from the chaos decomposition theorem.

Let us prove i). We will prove the formulas $a), b), c), d)$ for functions of the form $f_{k,q}$. Then the density of $P_{k,q}(\Delta_n), q \in \Theta^n_k, n \in N$ in H_k and of $f_{k,q}, q \in \Theta^n_k, n \in N$, in $L^2_{s,k}$ imply that $I_k(f_k) \in D^\infty \cap Dom(L)$, and $(I_{k-1}(f_k(s,\circ)))_{s\geq 0} \in Dom(\delta)$. The proof of $a), b), c), d)$ may be done starting with the explicit form of the operators D, δ, L on simple functionals, but this is a rather unpleasant computation. A more elegant way of doing it is to use the properties of these operators. Let us prove $a)$ in the case $k=2$ (the proof is analogues for general k and the proof of $b)$ is analogues also, so we live them out). Using (33) we write

$$\begin{aligned} D_s \int_0^1 \int_0^{t_1} f_{k,q}(t_1,t_2) dB_{t_2} dB_{t_1} &= \int_0^s f_{k,q}(s,t_2) dB_{t_2} + \int_s^1 D_s \int_0^{t_1} f_{k,q}(t_1,t_2) dB_{t_2} dB_{t_1} \\ &= \int_0^s f_{k,q}(s,t_2) dB_{t_2} + \int_s^1 f_{k,q}(t_1,s) dB_{t_1} \\ &= \int_0^1 f_{k,q}(s,t_2) dB_{t_2} = I_1(f_{k,q}(s,\circ). \end{aligned}$$

In order to prove $c)$ we use the duality between D and δ. For every $g_{k,q}$ we have

$$\begin{aligned} & E(\delta(I_{k-1}(f_{k,q}(s,\circ)))I_k(g_{k,q})) \\ = \; & E\int_0^1 I_{k-1}(f_{k,q}(s,\circ)) D_s I_k(g_{k,q}) ds = E\int_0^1 I_{k-1}(f_{k,q}(s,\circ)) I_{k-1}(g_{k,q}(s,\circ)) ds \\ = \; & \int_0^1 \ldots \int_0^1 f_{k,q}(t_1,\ldots,t_k) g_{k,q}(t_1,\ldots,t_k) dt_k \ldots dt_1 = E(I_k(f_{k,q}) I_k(g_{k,q})) \end{aligned}$$

which implyes $c)$. Let us prove $d)$ for $f_{k,q}$ (for a general f_k one procedes by approximation). This is an easy consequence of (32) and of the following simple facts: $L(\Delta^k_n) = -\Delta^k_n$ and $< D\Delta^k_n, D\Delta^p_n >= 0$ is $k \neq p$. *square*

Clearly the above formulas hold true for finite sums of multiple integrals. Then the question is if we are able to pass to the limit and to extend the above formulas to general series. The answer is given by the following theorem which give the characterizations of the domains of the operators in Malliavin calculus in terms of Weiner chaos expansions.

Theorem 5 *Let $f_k \in L^2_{s,k}, k \in N$ be the kernels in the chaos decomposition of $F \in L^2(\Omega)$.*
i) $F \in D^{1,2}$ if and only if

$$\sum_{k=0}^{\infty} k \times k! \, \|f_k\|^2_{L^2[0,1]^k} < \infty. \tag{52}$$

In this case

$$D_s F = \sum_{k=0}^{\infty} (k+1) I_k(f_{k+1}(s, \circ)) \quad \textit{and} \quad \|DF\|^2_{L^2([0,1]\times\Omega)} = \sum_{k=0}^{\infty} k \times k! \, \|f_k\|^2_{L^2[0,1]^k}. \tag{53}$$

ii) $F \in D^{p,2}$ if and only if

$$\sum_{k=0}^{\infty} ((k+p-1) \times ... \times k) \times k! \, \|f_k\|^2_{L^2[0,1]^k} < \infty. \tag{54}$$

In this case

$$D^{(p)}_{s_1,\ldots,s_p} F = \sum_{k=0}^{\infty} I_k(f_{k+p}(s_1, ..., s_p, \circ)) \quad \textit{and} \tag{55}$$

$$\left\| D^{(p)} F \right\|^2_{L^2([0,1]^p \times \Omega)} = \sum_{k=0}^{\infty} ((k+p-1) \times ... \times k) \times k! \, \|f_k\|^2_{L^2[0,1]^k}. \tag{56}$$

iii) $F \in Dom(L)$ if and only if

$$\sum_{k=0}^{\infty} k^2 \times k! \, \|f_k\|^2_{L^2[0,1]^k} < \infty. \tag{57}$$

and in this case

$$LF = -\sum_{k=1} k I_k(f_k). \tag{58}$$

Proof. i) It is easy to check that, if $F_n \in D^{1,2}, n \in N$ and $F_n \to F$ in $L^2(\Omega)$ and $DF_n \to G$ in $L^2([0,1] \times \Omega)$ then $F \in D^{1,2}$ and $DF = G$. Take $F_n = \sum_{k=0}^{n} I_k(f_k)$. The condition (52) says that the sequence $(DF_n)_{n \in N}$, is Cauchy in $L^2([0,1] \times \Omega)$ so guarantees that $F \in D^{1,2}$ and gives the expression of DF in (53). Let us now prove that if $F \in D^{1,2}$ then (52) holds true. Let V be the subspace of $L^2(\Omega)$ of the random variables for which (52) holds true. We already know that this is a subspace of $D^{1,2}$ and we want to prove that in fact this is the whole $D^{1,2}$. Note first that V is closed in $D^{1,2}$. In fact, if $F_n \in V, n \in N$ converges to F in $\|\circ\|_2$ then the kernels in the chaos decomposition converge and so, for every $m \in N$

$$\sum_{k=0}^{m} k \times k! \, \|f_k\|^2_{L^2[0,1]^k} = \lim_n \sum_{k=0}^{m} k \times k! \, \|f_k^n\|^2_{L^2[0,1]^k} \leq \sup_n \|DF_n\|^2_{L^2([0,1]\times\Omega)} =: C.$$

If $F_n, n \in N$ is Cauchy in $\|\circ\|_{1,2}$ then it is bounded and so $C < \infty$. Taking the $\sup_m$ we see that the series is finite and so $F \in V$. So V is closed and it remains to prove that if $F \in D^{1,2}$ is orthogonal to V with respect to the scalar product of $D^{1,2}$ then F is null. We take a polynomial $P_{k,q}$ and we recall that $LP_{k,q}(\Delta_n) = -kP_{k,q}(\Delta_n)$. Using the integration by parts formula we obtain

$$\begin{aligned} 0 &= E(FP_{k,q}(\Delta_n)) + E\int_0^1 D_sFD_sP_{k,q}(\Delta_n)ds \\ &= E(F(P_{k,q}(\Delta_n) - LP_{k,q}(\Delta_n)) = (1+k)E(FP_{k,q}(\Delta_n)). \end{aligned}$$

Since the above polynomials are dense in $L^2(\Omega)$ (see the previous lemma) this implies that $F = 0$.

ii) The proof of $ii)$ is quait analogues except for the orthogonality argument which is has to be adapted. Let $p > 1$ and let V_p be the subspace of $L^2(\Omega)$ of the random variables for which (54) holds true. As above, V_p is a closed linear subspace of $D^{p,2}$ which is a Hilbert space with the scalar product defined by

$$< F, G >_p =: E(FG) + \sum_{l=1}^{p} E < D^{(l)}F, D^{(l)}G >_{L^2[0,1]^l} .$$

Here $D^{(l)}F$ designes the derivatives of order l (and not the derivative with respect to B^l which was denoted by D^lF; anyway we have here only one Brownian mùotion). In order to prove that $V_l = D^{p,2}$ we have to check that if $F \in D^{p,2}$ is orthogonal on V_l then $F = 0$. We first note that, if G is a simple functional then $E < D^{(l)}F, D^{(l)}G >_{L^2[0,1]^l} = (-1)^l E(FL^lG)$. This holds true if F is a simple functional also (direct computation or iteration of the duality formula for $l = 1$) and then extends to $F \in D^{l,2}$ by approximation with simple functionals. Once this formula is known the reasoning is as above: we take some polynomial $P_{k,q}(\Delta_n)$ and recall that $L^lP_{k,q}(\Delta_n) = (-k)^lP_{k,q}(\Delta_n)$ and obtain

$$\begin{aligned} 0 &= < F, P_{k,q}(\Delta_n) >_p = E(FP_{k,q}(\Delta_n)) + \sum_{l=1}^{p} E < D^l F, D^l P_{k,q}(\Delta_n) >_{L^2[0,1]^l} \\ &= E(FP_{k,q}(\Delta_n)) \sum_{l=0}^{p} k^l \end{aligned}$$

so that $E(FP_{k,q}(\Delta_n))$ which implies $F = 0$.

iii). If $F \in Dom(L)$ then $LF \in L^2(\Omega)$ and so admits a decomposition in Wiener chaos with some kernels f_k^L. Let us check that $f_k^L = kf_k$. Once we have proved this equality (57) follows from the fact that $\|LF\|_2 < \infty$ and (58) is obvious... We take some arbitrary g_k and write

$$< g_k, f_k^L >_{L^2[0,1]^k} = E(I_k(g_k)LF) = E(LI_k(g_k)F) = kE(I_k(g_k)F) = k < g_k, f_k >_{L^2[0,1]^k} .$$

This clearly implies $f_k^L = kf_k$ and the proof is completed. □

Remark 7 *The Weiner chaos expansion theorem says that F is in $L^2(\Omega)$ if the series $\sum_{k=0}^{\infty} k! \|f_k\|^2_{L^2[0,1]^k}$ converges. This theorem says that the "regularity" of F depends on the speed of convergence of this series and gives a precise description of this phenomenons. If one looks to the multiple integrals $I_k(f_k)$ as to the "smooth" functionals (the analogues of C^∞ functions) then the fact that the series converges very fast means that the functional is closed to the partial series and so to the smooth functionals. This also suggests to take the multiple integrals as the basic objects (replacing the simple functionals in our approach), to define the operators in Malliavin calculus by means of (51) and then to extend these operators to their natural domain given in the previous theorem. This is the approach in [16].*

Remark 8 *As an immediate consequence of (55) one obtains the so called Stroock formula: if $F \in D^{p,2}$ then*

$$f_p(s_1, ..., s_p) = \frac{1}{p!} E(D^{(p)}_{s_1,...,s_p} F) \tag{59}$$

which may be very useful in order to control the chaos kernels f_k.

Remark 9 *The operator L is the infinitesimal operator of the Ornstein Uhlembeck semigroup $T_t : L^2(\Omega) \to L^2(\Omega)$ defined by $T_t F = \sum_{k=0}^{\infty} e^{-kt} I_k(f_k)$. See [15] for this topic.*

Remark 10 *One defines the operator $C = -\sqrt{-L}$ (so $C(I_k(f_k)) = -\sqrt{k} I_k(f_k)$). Note that $D^{1,2} = Dom(C)$. The following inequalities due to Meyer [13] (and known as Meyer's inequalities) are central in the analysis on the Weiner space and in the distribution theory on the Weiner space due to Watanabe (see [20] and [10]) . For every $p > 1$ and every $k \in N$ there are some constants $c_{p,k}$ and $C_{p,k}$ such that, for every $F \in D^\infty$*

$$c_{p,k}E\left\|D^{(k)}F\right\|_{L^2[0,1]^k}^p \leq E\left|C^kF\right|^p \leq C_{p,k}(E\left|F\right|^p + E\left\|D^{(k)}F\right\|_{L^2[0,1]^k}^p)$$

See [15] for this topic.

We close this section with a relative compactness criterion on the Wiener space (it was first been proved in [6] and then a version in Sobolev-Wiener spaces has been given in [1, 17]).

Theorem 6 *Let $F_n, n \in N$ be a bounded sequence in $D^{1,2}$ (that is $\sup_n \|F_n\|_{1,2} < \infty$). Suppose that*

i) There is a constant C such that for every $h > 0$

$$\sup_n E\int_0^{1-h} \left|D_{t+h}F_n - D_tF_n\right|^2 dt \leq Ch. \tag{60}$$

ii) For every $\varepsilon > 0$ there is some $\eta > 0$ such that

$$\sup_n E\int_{[0,\eta)\cup(1-\eta,1]} \left|D_tF_n\right|^2 dt \leq \varepsilon. \tag{61}$$

Then $(F_n)_{n\in N}$ is relatively compact in $L^2(\Omega)$.

Proof. Step 1. Let $F_n = \sum_{k=0}^{\infty} I_k(f_k^n)$.
For every integer N

$$E\left|\sum_{k=N}^{\infty} I_k(f_k^n)\right|^2 = \sum_{k=N}^{\infty} k!\,\|f_k^n\|_{L^2[0,1]^k}^2 \leq \frac{1}{N}\sum_{k=N}^{\infty} k{\times}k!\,\|f_k^n\|_{L^2[0,1]^k}^2 \leq \frac{1}{N}\sup_n E\int_0^1 \left|D_sF_n\right|^2 ds.$$

So it will suffice to prove that for each fixed k the sequence $I_k(f_k^n), n \in N$ is relatively compact in $L^2(\Omega)$ and this is equivalent to the relative compactness of the sequence $f_k^n, n \in N$, in $L^2[0,1]^k$ (one employes the isometry property).

Step 2. We will prove that for each $h_1, ..., h_k > 0$

$$\int_0^{1-h_1} ... \int_0^{1-h_k} \left|f_k^n(t_1 + h_1, ..., t_k + h_k) - f_k^n(t_1, ..., t_k)\right|^2 dt_1...dt_k \leq C(h_1 + ... + h_k) \tag{62}$$

where C is a constant independent of n and on $h_1, ..., h_k$.

Moreover, for each $\varepsilon > 0$ one may find $\eta > 0$ such that

$$\sup_n \int_{D_\eta} \left|f_k^n(t_1, ..., t_k)\right|^2 dt \leq \varepsilon. \tag{63}$$

where $D_\eta = \{(t_1, ..., t_k) : t_i \in [0,\eta) \cup (1-\eta, 1]$ for some $i = 1, ..., k\}$. Once these two properties proved, a classical relative compactness criterion in $L^2[0,1]^k$ guarantees that the sequence $f_k^n, n \in N$, is relatively compact.

Since f_k^n is symmetric (62) reduces to

$$\int_0^{1-h_1}\int_0^1 ... \int_0^1 \left|f_k^n(t_1+h_1,t_2,....,t_k)-f_k^n(t_1,...,t_k)\right|^2 dt_1...dt_k \leq Ch_1.$$

For each fixed $t_1 \in [0,1]$

$$\begin{aligned}
& \int_0^1 ... \int_0^1 \left|f_k^n(t_1+h_1,t_2,....,t_k)-f_k^n(t_1,...,t_k)\right|^2 dt_2...dt_k \\
= \quad & E\left|I_{k-1}(f_k^n(t_1+h_1,\circ))-I_k(f_k^n(t_1,\circ))\right|^2 \leq \sum_{p=1}^{\infty} E\left|I_{p-1}(f_p^n(t_1+h_1,\circ))-I_{p-1}(f_p^n(t_1,\circ))\right|^2 \\
= \quad & E\left|D_{t_1+h_1}F_n - D_{t_1}F_n\right|^2 \leq Ch_1
\end{aligned}$$

and (62) is proved. The proof of (63) is analogues. □

4 Abstract Malliavin Calculus in finite dimension

The specificity of the Malliavin calculus is to be a differential calculus in infinite dimesion, so it seems strange that we begin by presenting it in finite dimension. We do this for two reasons. First of all in order to understand well which are the facts which permite to pass from finite dimension to infinite dimension. Moreover, this calculus was concieved in the "Gaussian frame" and then extends to Poisson processes as well. Warking with such lows permits to pass from the finite dimenson to the infinite dimension. But the operators introduced in Malliavin calculus and the strategy used in order to obtain an integration by parts formula remain interesting even in the finite dimensional case - the fact that these operatros and formulas pass to the limit is some how a test which show that they are the good objects. This is why there is some interest in following the same strategy in order to obtain integration by parts formulas for a ruther large class of random variables which go far beond Gaussian or Poisson. But we do no more expect these formulas to pass to the limit in the general case.

In this section there are two objects which are given:

◇ An $m-$dimensional random variable

$$H=(H_1,...,H_m).$$

We assume that the law of H has a density p_H with respect to the Lebesgue measure and this density is one time differentiable. We denote

$$\theta_i = \frac{\partial \ln p_H}{\partial x_i} = \frac{1}{p_h} \times \frac{\partial p_h}{\partial x_i}.$$

◇ A poisitve measure $\mu = (\mu_1,...,\mu_m)$ on $\{1,...,m\}$.

The random variable H is the one which naturaly appears in our problem. The choice of the measure μ is more problematic - but here we assume that it is given.

Simple functionals and simple processes. We denote by S_m the space of the "simple functionals" $F = f(H)$ where $f \in C^\infty(R^m, R)$ and has polynomial growth and by P_m the space of the "simple processes" $U = (U_1, ..., U_m)$ where $U_i = u_i(H)$ are simple functionals. We think to S_m as to a subspace of $L^2(\Omega, F, P)$ where (Ω, F, P) is the probability space on which H is defined. And we think to P_m as to a subspace of $L^2(\Omega \times \{1, ..., m\}, F \otimes T, P \otimes \mu)$ where T designes the σ field of all the sub-sets of $\{1, ..., m\}$.

Didfferential operators. We define a "Malliavin's derivative operator" $D : S_m \to P_m$ and an "Skorohod integral operator" $\delta : P_m \to S_m$ in the following way. Following the Malliavin calculus we put $DF = U$ with $U_i = \frac{\partial f}{\partial x_i}(H)$. We denote $D_i F = U_i = \frac{\partial f}{\partial x_i}(H)$. We look now for δ. We construct it as the adjoint of D in the L^2 spaces mentioned above. Let $U = (U_1, ..., U_m)$ with $U_i = u_i(H)$ and $F = f(H)$. We write

$$\begin{aligned} E\left(\sum_{i=1}^m \mu_i D_i F \times U_i\right) &= E\left(\sum_{i=1}^m \mu_i \frac{\partial f}{\partial x_i}(H) \times u_i(H)\right) = \sum_{i=1}^m \mu_i \int \frac{\partial f}{\partial x_i}(y) \times u_i(y) p_H(y) dy \\ &= -\sum_{i=1}^m \mu_i \int f(y) \times \left(\frac{\partial u_i}{\partial x_i}(y) p_H(y) + u_i(y) \frac{\partial p_H}{\partial x_i}(y)\right) dy \\ &= -\int f(y) \times \left(\sum_{i=1}^m \mu_i \left(\frac{\partial u_i}{\partial x_i}(y) + u_i(y) \frac{\partial \ln p_H}{\partial x_i}(y)\right)\right) p_H(y) dy \\ &= -E(F\left(\sum_{i=1}^m \mu_i \left(\frac{\partial u_i}{\partial x_i}(H) + u_i(H)\theta_i(H)\right)\right)) = E(F\delta(U)) \end{aligned}$$

with

$$\delta(U) =: -\sum_{i=1}^m \mu_i \left(\frac{\partial u_i}{\partial x_i}(H) + u_i(H)\theta_i(H)\right).$$

For $U, V \in P_m$ we introduce the scalar product

$$< U, V >_\mu = \sum_{i=1}^m (U_i \times V_i)\, \mu_i.$$

With this notation the above euqlity reads

$$E(< DF, U >_\mu = E(F\delta(U)).$$

Finally we define the corresponding "Orenstein Ulemback operator" $L : S_m \to S_m$ as $LF =: \delta(DF)$. This gives

$$L(F) = -\sum_{i=1}^m \mu_i \left(\frac{\partial^2 f}{\partial x_i^2}(H) + \frac{\partial f}{\partial x_i}(H)\theta_i(H)\right)$$

The standard rules of the differential calculus give the following chain rule.

Proposition 10 *Let* $F=(F^1,...,F^d)$ $F^1,...,F^d\in S_m$ *and* $\phi: R^d\to R$ *is a smooth function. Then* $\phi(F)\in S_m$ *and*

$$i)\quad D_i\phi(F)=\sum_{j=1}^m \frac{\partial\phi}{\partial x^j}(F)D_iF^j$$

$$ii)\quad L\phi(F)=-\sum_{i,j=1}^m \frac{\partial^2\phi}{\partial x^i\partial x^j}(F)<DF^j,DF^j>_\mu+\sum_{j=1}^m\frac{\partial\phi}{\partial x^j}(F)LF^j.$$

In particular, taking $\phi(x^1,x^2)=x^1\times x^2$ *we obtain*

$$iii)\quad L(FG)=FLG+GLF-2<DF,DG>_\mu .$$

We are now ready to derive integration by parts formulas of type $IP_\alpha(F;G)$. So we look to

$$E(\frac{\partial\phi}{\partial x_i}(F)G)$$

where $\phi: R^d\to R$ is a smooth function, $F=(F^1,...,F^d)$ and $F^1,...,F^d,G\in S_m$. We define σ_F to be the analogues of Malliavin's covarience matrix, that is

$$\sigma_F^{ij}=:<DF^i,DF^j>_\mu=\sum_{r=1}^m(\frac{\partial f^i}{\partial x_r}\frac{\partial f^j}{\partial x_r})(H)\mu_r$$

where $F^i=f^i(H)$. We suppose that σ_F is invertible and we denote $\gamma_F=:\sigma_F^{-1}$. Using the chain rule we write

$$\begin{aligned}<\quad & D\phi(F),DF^j>_\mu=\sum_{r=1}^m D_r\phi(F)D_rF^j\mu_r=\sum_{r=1}^m\left(\sum_{q=1}^m\frac{\partial\phi}{\partial x^q}(F)D_rF^q\right)D_rF^j\mu_r\\ = \quad & \sum_{q=1}^m\frac{\partial\phi}{\partial x^q}(F)\sum_{r=1}^m D_rF^qD_rF^j\mu_r=\sum_{q=1}^m\frac{\partial\phi}{\partial x^q}(F)<DF^q,DF^j>_\mu=\sum_{q=1}^m\frac{\partial\phi}{\partial x^q}(F)\sigma_F^{qj}.\end{aligned}$$

It follows that

$$\frac{\partial\phi}{\partial x^q}(F)=\sum_{j=1}^m<D\phi(F),DF^j>_\mu\ \gamma_F^{jq}.$$

Moreover we have

$$-2<D\phi(F),DF^j>_\mu=L(\phi(F)F^j)-\phi(F)LF^j-F^jL\phi(F)$$

so that

$$\begin{aligned}
E(\frac{\partial\phi}{\partial x_i}(F)G) &= E(\sum_{j=1}^{m} < D\phi(F), DF^j > \gamma_F^{ji}G) \\
&= \frac{1}{2}E(\sum_{j=1}^{m}\left(\phi(F)LF^j + F^jL\phi(F) - L(\phi(F)F^j)\right)\gamma_F^{ji}G) \\
&= \frac{1}{2}E(\phi(F)\sum_{j=1}^{m}\Big(\gamma_F^{ji}GLF^j + L(F^j\gamma_F^{ji}G) - F^jL(\gamma_F^{ji}G)\Big)).
\end{aligned}$$

So we have proved the following integration by parts formula.

Theorem 7 *Let $F = (F^1, ..., F^d)$ and G, be such that $F^j, G \in S_m$. Suppose that σ_F is invertible and the inverse* $\det\gamma_F$ *has finite moments of any order. Then for every smooth ϕ*

$$E(\frac{\partial\phi}{\partial x_i}(F)G) = E(\phi(F)H^i(F;G))$$

with

$$H^i(F;G) = \frac{1}{2}\sum_{j=1}^{m}\Big(\gamma_F^{ji}GLF^j + L(F^j\gamma_F^{ji}G) - F^jL(\gamma_F^{ji}G)\Big).$$

This is the central integration by parts formula in Malliavin's calculus (in finite dimensional version).

The non-degenrency problem. We need $H^i(F;G)$ to be at list integrable (see the previous section about the density of the law of F and about conditional expectations). Since F and G are under control the only problem is about γ_F. So we have to study the property

$$E\left|\det\gamma_F\right|^p < \infty.$$

This is related to $\lambda_F > 0$ where λ_F is the smaller proper value of σ_F which is computed as

$$\begin{aligned}
\lambda_F &= \inf_{|\xi|=1} < \sigma_F\xi, \xi >= \inf_{|\xi|=1}\sum_{r=1}^{m}\mu_r\sum_{i,j=1}^{d}\frac{\partial f^i}{\partial x_r}(H)\frac{\partial f^j}{\partial x_r}(H)\xi^i\xi^j \\
&= \inf_{|\xi|=1}\sum_{r=1}^{m}\mu_r\left(\sum_{i=1}^{d}\frac{\partial f^i}{\partial x_r}(H)\xi^i\right)^2.
\end{aligned}$$

Recall that

$$\det\sigma_F \geq \lambda_F^d$$

We define the real function

$$\Gamma_F(x) = \inf_{|\xi|=1}\sum_{r=1}^{m}\mu_r\left(\sum_{i=1}^{d}\frac{\partial f^i}{\partial x_r}(x)\xi^i\right)^2$$

Since $\lambda_F = \Gamma_F(H)$, the function Γ_F control the non degenerency of σ_F. Note that $\Gamma_F(x) > 0$ means that $Span\{v_r(x), r = 1, ..., m\} = R^d$ where $v_r = (\frac{\partial f^1}{\partial x_r}, ..., \frac{\partial f^d}{\partial x_r})$. So it is the analoguoues of the ellipticity condition in the frame of diffusion processes. Note also that, if we suppose that $f^i, i = 1, ..., d$ are continuously differentiable (and we does!) then $\Gamma_F(x_0) > 0$ for some point x_0 implies that $\Gamma_F(x)$ is minorated on a whole neighbourhod of x_0. This suggests that at list in a first etape we may evacuate the difficulty of the non degenerency problem just by using local results with a localization in the region where $\Gamma_F(x)$ is strictly positive. We have the following obvious fact:

Lemma 8 *Let $D \subset R^d$ be a set such that $\Gamma_F(x) \geq c > 0$ for every $x \in D$ and let $\eta \in C^\infty(R^d; R)$ with $\sup \eta \subset D$. Then*

$$E\frac{\eta(H)}{|\det \sigma_F(H)|^p} < \infty, \quad \forall p \in N.$$

This is because $\eta(H) \neq 0 \Rightarrow \det \sigma_F(H) \geq \lambda_F^d = \Gamma_F(H)^d \geq c^d$.

This suggests the following localized version of the integration by parts formula.

Theorem 8 *We assume the same hypothesis as in the previous theorem - except the non degenerency - and we consider a function $\eta \in C^\infty(R^d; R)$ with $\sup \eta \subset \{\Gamma_F > c\}$ for some $c > 0$. Then*

$$IP_\eta^i(F; G) \quad E(\frac{\partial \phi}{\partial x_i}(F)\eta(H)G) = E(\phi(F)H_\eta^i(F; G))$$

with

$$H_\eta^i(F; G)) = \frac{1}{2}\sum_{j=1}^m \left(\gamma_F^{ij}\eta(H)GLF^j + L(F^j\eta(H)\gamma_F^{ij}G) - F^jL(\eta(H)\gamma_F^{ij}G)\right).$$

Density of the law. A first application of this obvious localization procedure is that we are able to obtain local densities for tha law of F. More precisely we say that F has a density p on an open domain $D \subset R^d$ if for every $\phi \in C^\infty(R^d; R)$ with $\sup \phi \subset D$ one has

$$E(\phi(F)) = \int \phi(x)p(x)dx.$$

We have the following standard result (we give first the result in the one dimensional case):

Lemma 9 *Let $F \in S_m$. The law of F has a density on $\{\Gamma_F > 0\}$. An integral representation of the density p may be obtained in the following way. Let $c > 0$ and let $\eta \in C^\infty(R^d; R)$ with $\eta(x) = 1$ for $x \in \{\Gamma_F > c/2\}$ and $\eta(x) = 0$ for $x \in \{\Gamma_F < c/4\}$. Then for every $x \in \{\Gamma_F > c/2\}$*

$$p(x) = E(1_{[x,\infty)}(F)H_\eta(F; 1)). \tag{64}$$

Proof. The formal argument is the following: since $\delta_0(y) = \partial_y 1_{[0,\infty)}(y)$ and $\theta(y) = 1$ on a neighbourhod of x, one employes $IP_\eta(F;1)$ in order to obtain

$$\begin{aligned}E(\delta_0(F-x)) &= E(\partial_y 1_{[0,\infty)}(F-x)) = E(\partial_y 1_{[0,\infty)}(F-x)\eta(y)) \\ &= E(1_{[0,\infty)}(F-x)H_\eta(F;1)) = E(1_{[x,\infty)}(F)H_\eta(F;1)).\end{aligned}$$

In order to let this reasoning rigourous one has to regularize the Dirac function. So we take a positive function $\phi \in C_c^\infty(R)$ with the support equal to $[-1,1]$ and such that $\int \phi(y)dy = 1$ and for each $\delta > 0$ we define $\phi_\delta(y) = \delta^{-1}\phi(y\delta^{-1})$. Moreover we define Φ_δ to be the primitive of ϕ_δ given by $\Phi_\delta(y) = \int_{-\infty}^y \phi_\delta(z)dz$ and we construct some random variables θ_δ of law $\phi_\delta(y)dy$ and which are independent of F. For each $f \in C_c^\infty(R)$ we have

$$E(f(F)) = \lim_{\delta\to 0} Ef(F-\theta_\delta)). \tag{65}$$

Supose now that $\sup f \subset \{\Gamma > c\}$ and take the function η from the ennonce. We write

$$Ef(F-\theta_\delta)) = \int\int f(u-v)\phi_\delta(v)dvdP\circ F^{-1}(u) = \int\int f(z)\phi_\delta(u-z)dzdP\circ F^{-1}(u).$$

Suppose that δ is sufficiently small in order that $z \in \{\Gamma > c\}$ and $|u-z| < \delta$ implies $\Gamma(u) > c/2$. Then $\phi_\delta(u-z) = \phi_\delta(u-z)\eta(u)$ so we obtain

$$\begin{aligned}\int\int f(z)\phi_\delta(u-z)dzdP\circ F^{-1}(u) &= \int\int f(z)\phi_\delta(u-z)\eta(u)dzdP\circ F^{-1}(u) \\ &= \int f(z)E(\phi_\delta(F-z)\eta(F))dz = \int f(z)E(\Phi_\delta'(F-z)\eta(F))dz \\ &= \int f(z)E(\Phi_\delta(F-z)H_\eta(F;1))dz.\end{aligned}$$

The above relation togather with (65) guarantees that the law of F is absolutely continuous with respect to the Lebesgue measure. On the other hand $\Phi_\delta(y) \to 1_{[x,\infty)}(y)$ except for $y = 0$, so $\Phi_\delta(F-z) \to 1_{[0,\infty)}(F-z) = 1_{[z,\infty)}(F), P-a.s.$ Then using Lebesgues dominated convergence theorem we pass to the limit in the above relation and we obtain

$$E(f(F)) = \int f(z)E(1_{[z,\infty)}(F)H_\eta(F;1))dz.$$

□

Let us now give the multi dimensional version of the above result. We consider a multi-index $\alpha = (\alpha_1,...,\alpha_q) \in \{1,...,d\}^q$, we denote $|\alpha| = q$ and $D^\alpha = \partial^q/\partial x_{\alpha_1}...\partial x_{\alpha_q}$. We also define by reccurence $H_\theta^\alpha(F) = H_\theta^{\alpha_q}(F; H_\theta^{\alpha'}(F))$ where $\alpha' = \{\alpha_1,...,\alpha_{q-1}\}$. In order to be able to define these quantities we need the non degenerency of σ_F, but, if we suppose that $\sup\eta \subset \{\Gamma_F > 0\}$ we have already seen that this property holds true.

Lemma 10 *i) Let $F \in S_m^d$. The law of F has a density on $\{\Gamma_F > 0\}$. An integral representation of p may be obtained in the following way. Let $c > 0$ and let $\eta \in C^\infty(R^d; R)$ with $\eta(x) = 1$ for $x \in \{\Gamma > c/2\}$ and $\eta(x) = 0$ for $x \in \{\Gamma_F < c/4\}$. Then for every $x \in \{\Gamma_F > c\}$*

$$p(x) = E(1_{I(x)}(F)H_\eta^{(1,\ldots,1)}(F)).$$

ii) Suppose that $\Gamma_F(x) > 0$ for every $x \in R^d$. Then F has a density on the whole R^d.

The first point is proven as in the one dimensional case, just noting that $D^{(1,\ldots,d)}1_{I(x)}(y) = \delta_x(y)$. The second point is a trivial consequence of the first one: if one has a local density in each reagion, of course one has a global density. There is here a point which seems ruther amaising because usually we need "global" non degenerency for the Malliavin matrix in order to produce a density. Or, in order to obtain "global non degenerency" (this means $E \det \sigma_F^{-k} < \infty$ for every k) we need $\Gamma_F(x) > c$ for every x, which amounts to an uniform ellipticity assumtion. And such an assumption is more restrictive then ours, because one may have $\lim_{x\to\infty} \Gamma(x) = 0$. In fact, as it is clear from our reasoning, the uniform ellipticity is not necessary in order to produce a density, and moreover, in order to prove that this density is smooth. At the contrary, if we want to obtain some evaluations concerning the behaviour of this density as $x \to \infty$ then the localization argument does no more work and we need the uniform ellipticity assumption. This is moral.

The sensitivity problem. We assume that the random variable F is now indiciated on some $\rho \in U$ where U is an open domain in some R^k. So we have $F^i(\rho, \omega) = f^i(\rho, H), i = 1, ..., d$ where f^i are some smooth functions, both of ρ and of x. We would consider that H depends also on the parameter ρ but this is a little bit more complex situation which we live out for the moment - we assume that we have succeded to emphasise some 'objective' random variable H on which the model depends. We are interested in the derivatives with respect to ρ of $J(\rho) =: E(\phi(F(\rho, \omega)))$ where ϕ is some function. Suppose that ϕ is differentiable and we may comute the derivatives with the expectations. Then

$$\frac{\partial J}{\partial \rho_i}(\rho) = \sum_{j=1}^{d} E(\frac{\partial \phi}{\partial x_j}(F(\rho,\omega))\frac{\partial f^j}{\partial \rho_i}(F(\rho,\omega)).$$

If ϕ is differentiable we stop here and we use the above formula in order to compute the derivatives of J. But, if ϕ is not differentiable, then we have to use the integration by parts in order to obtain

$$\frac{\partial J}{\partial \rho_i}(\rho) = E(\phi(F(\rho,\omega))\sum_{j=1}^{d} H^j(F(\rho,.); \frac{\partial f^j}{\partial \rho_i}(F(\rho,\omega))).$$

And this is the ibntegral representation of $\frac{\partial J}{\partial \rho_i}$.

What about non degenrency and about localization? The non degenerency set is $\{\Gamma_\rho > 0\}$ with

$$\Gamma_\rho(x) = \inf_{|\xi|=1} \sum_{r=1}^{m} \mu_r \left(\sum_{i=1}^{d} \frac{\partial f^i}{\partial x_r}(\rho, x)\xi^i \right)^2$$

If we localize on this set we need no more non- degnerency assumption (of course, if $\Gamma_\rho > 0$ everywere, there is no problem). Now, what we hope is that the singularities of ϕ are in the region of non degenerancy. In order to write it down we denote by $\Lambda_{\rho_0,\varepsilon,c} = \{x : \Gamma_\rho(f(\rho,x)) > c, \forall\rho \in (\rho_0-\varepsilon,\rho_0+\varepsilon)\}$ and we consider a localization function π which is equal to 1 on $\Lambda_{\rho_0,\varepsilon,c}$ and si null outside $\Lambda_{\rho_0,\varepsilon/2,c/2}$. We also suppose that ϕ is differentiable outside $\Lambda_{\rho_0,\varepsilon,c}$. Then we write

$$\begin{aligned}\frac{\partial J}{\partial\rho_i}(\rho) &= \sum_{j=1}^{d} E(\frac{\partial}{\partial x_j}(\phi\pi)+\phi(1-\pi))(F(\rho,\omega))\frac{\partial f^j}{\partial\rho_i}(F(\rho,\omega)) \\ &= \sum_{j=1}^{d} E(\frac{\partial\phi(1-\pi)}{\partial x_j}(F(\rho,\omega))\frac{\partial f^j}{\partial\rho_i}(F(\rho,\omega)) \\ &\quad +E(\phi(F(\rho,\omega))\sum_{j=1}^{d} H^i(F(\rho,.);\frac{\partial f^j}{\partial\rho_i}(F(\rho,\omega)\pi(F(\rho,\omega))).\end{aligned}$$

So we simplely derivate in the region in which ϕ is differentiable and we employ the integration by parts formula in the region where ϕ is not differentiable. If we are lucky the singularitites of ϕ ar in the non degenerency rigion of H (in $\Lambda_{e,\rho_0,c}$) and so the non degenerancy assumption is automatically satisfied.

5 Diffusion processes

In this section we briefly present the Malliavin calculus for diffusion process. We consider the $N-$dimensional diffusion process X solution of the SDE

$$dX_t^i = \sum_{j=1}^{d} \sigma_j^i(X_t)dB_t^j + b^i(X_t)dt, \quad i=1,N. \tag{66}$$

We denote by $C_{l,b}^\infty(R^N;R)$ the infinitely differentiable functions which have linear growth and have bounded derivatives of any order. We assume that

$$b^i,\sigma_j^i \in C_{l,b}^\infty(R^N;R), \quad i=1,...,N, j=1,...,d. \tag{67}$$

We denote by $X_t^x = (X_t^{x,1},...,X_t^{x,N})$ the solution starting from $x, i.e. X_0^x = x$.

The first application of Malliavin calculus was to prove that under Hormander's condition (see (H^{x_0}) bellow) the law of X_t has a smooth density with respect to the Lebesgue measure and to obtain exponential bounds for the density and for its derivatives (it turns out that this produces a probabilistic proof for Hormander's theorem - see [20] or [15]). Malliavin's approach goes through the absolute continuity criterion presented in the previous section. In order to apply this criterion one has to prove two things. First of all one has to check the regularity of X_t in Malliavin's sense, that is $X_t \in D^\infty$.

This is a long but straightforward computation. Then one has to prove that under Hormander's condition X_t is non degenerated, that is that (H_{X_t}) holds true. This is a problem of stochastic calculus which is much more difficult. A complete study of this problem (including very useful quantitative evaluations for the Malliavin's covariance matrix and for the density of the law of X_t) has been done in [12] (see also [15] or [10]). The subject is extremely technical so we do not give here complete proofs. We will give the main results and outline the proofs, just to give an idea about what is going on.

We begin with the results. Let us introduce the Lie bracket of $f, g \in C^1(R^N; R^N)$ defined by $[f, g] = f\nabla g - g\nabla f$ or, on components

$$[f, g]^i = \sum_{j=1}^{N} (\frac{\partial g^i}{\partial x^j} f^j - \frac{\partial f^i}{\partial x^j} g^j).$$

Then we construct by recurrence the sets of functions $L_0 = \{\sigma_1, ..., \sigma_d\}, L_{k+1} = \{[b, \phi], [\sigma_1, \phi], ..., [\sigma_d, \phi] : \phi \in L_k\}$ where σ_j is the $j'th$ column of the matrix σ. We also define $L_\infty = \cup_{k=0}^{\infty} L_k$. Given a fixed point $x_0 \in R^N$ we consider the hypothesis

$$(H_k^{x_0}) \quad Span\{\phi(x_0) : \phi \in \cup_{l=1}^{k} L_l\} = R^N.$$

We say that Hormander's hypothesis holds in x_0 if for some $k \in N$, (H_k^{xl0}) holds true, that is if

$$(H^{x_0}) \quad Span\{\phi(x_0) : \phi \in L_\infty\} = R^N.$$

Note that $(H_0^{x_0})$ means that $Span\{\sigma_1(x_0), ..., \sigma_d(x_0)\} = R^N$ which is equivalent with $\sigma\sigma^*(x_0) > 0$ which is the ellipticity assumption in x_0. So Hormander's condition is much weaker than the ellipticity assumption in the sense that except for the vectors $\sigma_1(x_0), ..., \sigma_d(x_0)$ it employes the Lie brackets as well.

Theorem 9 *i) Suppose that (67) holds true. Then for every $t \geq 0, X_t \in D^\infty$ and*

$$\|X_t^x\|_{k,p} \leq C_{k,p}(t)(1 + |x|)^{\beta_{k,p}} \tag{68}$$

where $C_{k,p}(t)$ is a constant which depends on k, p, t and on the bounds of b and σ and of their derivatives up to order k.

ii) Suppose that $(H_k^{x_0})$ holds true also. Then there are some constants $C_{k,p}(t), n_k \in N$ and $m_k \in N$ such that

$$\left\|(\det \gamma_{X_t^x})^{-1}\right\|_p \leq \frac{C_{k,p}(t)(1 + |x_0|)^{m_k}}{t^{n_k/2}}. \tag{69}$$

The function $t \to C_{k,p}(t)$ is increasing.. In particular the above quantity blows up as $t^{-n_k/2}$ as $t \to 0$.

iii) Suppose that (67) and (H^{x_0}) hold true. Then for every $t \geq 0$, the law of $X_t^{x_0}$ is absolutely continuous with respect to the Lebesgue measure and the density $y \to p_t(x_0, y)$ is a C^∞ function. Moreover, if σ and b are bounded, one has

$$\begin{aligned} p_t(x_0,y) &\leq \frac{C_0(t)(1+|x_0|)^{m_0}}{t^{n_0/2}} \exp(-\frac{D_0(t)\,|y-x_0|^2}{t}) \quad (70) \\ \left|D_y^{\alpha} p_t(x_0,y)\right| &\leq \frac{C_\alpha(t)(1+|x_0|)^{m_\alpha}}{t^{n_\alpha/2}} \exp(-\frac{D_\alpha(t)\,|y-x_0|^2}{t}) \end{aligned}$$

where all the above constants depend on the first level k for which $(H_k^{x_0})$ holds true and the functions C_0, D_0, C_α and D_α are increasing functions of t.

Remark 11 *In fact p is a smooth function of t and of x_0 as well and evaluations of the type (70) hold true for the derivatives with respect to these variables also. See [12] for complete information and proofs. Anyway it is alredy clear that the integral representation of the density given in (11) and in (12) give access to such informations.*

In the following we give the main steps of the proof.

The regularity problem. We sketch the proof of i). Let $n \in N$ and let $\overline{X}$ be the Euler scheme of step 2^{-n} defined by

$$\overline{X}^i(t_n^{k+1}) = \overline{X}^i(t_n^k) + \sum_{j=1}^{d} \sigma_j^i(\overline{X}(t_n^k))\Delta_n^{k,j} + b^i(\overline{X}(t_n^k))\frac{1}{2^n} \quad (71)$$

and $\overline{X}(0) = x$. We interpolate on $[t_n^k, t_n^{k+1})$ keeping the coefficients $\sigma_j^i(\overline{X}(t_n^k))$ and $b^i(\overline{X}(t_n^k))$ constants but we allow the Brownian motion and the time to move. This means that $\overline{X}(t)$ solves the SDE

$$\overline{X}(t) = x + \sum_{j=1}^{d} \int_0^t \sigma_j(\overline{X}(\tau_s))dB_s^j + \int_0^t b(\overline{X}(\tau_s))ds$$

where $\tau_s = t_n^k$ for $s \in [t_n^k, t_n^{k+1})$.

In view of (71) $\overline{X}(t_n^k)$ is a simple functional. So, in order to prove that $X_t \in D^{1,p}$, we have to prove that $\overline{X}(\tau_t)$ converges in $L^p(\Omega)$ to $X(t)$ - and this is a standard result concerning the Euler scheme approximation - and that $D\overline{X}(\tau_t)$ converges in $L^p([0,\infty) \times \Omega)$ to some limit. Then we define $DX(t)$ to be this limit. Using (33)

$$\begin{aligned} D_s^q \overline{X}^i(t) &= \sigma_q^i(\overline{X}(\tau_s)) + \sum_{j=1}^{d} \int_s^t \sum_{l=1}^{N} \frac{\partial \sigma_j^i}{\partial x^l}(\overline{X}(\tau_r)) D_s^q \overline{X}^i(\tau_r) dB_r^j \\ &\quad + \int_0^t \sum_{l=1}^{N} \frac{\partial b^i}{\partial x^l}(\overline{X}(\tau_r)) D_s^q \overline{X}^i(\tau_r) dr. \end{aligned}$$

We have used here the obvious fact that $D_s^q \overline{X}^i(\tau_r) = 0$ for $r < s$ which follows from the very definition of the Malliavin derivative of a simple functional. Assume now that s is fixed and let $Q_s(t), t \geq s$ be the solution of the $d \times N-$dimensional SDE

$$Q_s^{q,i}(t) = \sigma_q^i(X(s)) + \sum_{j=1}^{d} \int_s^t \sum_{l=1}^{N} \frac{\partial \sigma_j^i}{\partial x^l}(X(r)) Q_s^{q,l}(r) dB_r^j + \int_0^t \sum_{l=1}^{N} \frac{\partial b^i}{\partial x^l}(X(r)) Q_s^{q,l}(r) dr.$$

Then $D_s^q \overline{X}^i(t), t \geq s$ is the Euler scheme for $Q_s^{q,i}(t), t \geq s$ and so standard arguments give $\left\| D_s \overline{X}(t) - Q_s(t) \right\|_p \leq C_p 2^{-n/2}, \forall p > 1$ A quick inspection of the arguments leading to this inequality shows that C_p does not depend on s. Define now $Q_s(t) = D_s \overline{X}(t) = 0$ for $s \geq t$. We obtain

$$E\left(\int_0^\infty \left| Q_s(t) - D_s \overline{X}(t) \right|^2 ds \right)^{p/2} = E\left(\int_0^t \left| Q_s(t) - D_s \overline{X}(t) \right|^2 ds \right)^{p/2} \to 0$$

and so $X_t \in D^{1,p}$ and $D_s X(t) = Q_s(t)$. Recall that $D_. X(t)$ is an element of $L^2[0, \infty)$ and so is determined ds almost surely.

But we have here a precise version $Q_s(t)$ which is continuous and solves a SDE. So now on we will reefer to the Malliavin derivative of $X(t)$ as to the solution of

$$D_s^q X^i(t) = \sigma_q^i(X(s)) + \sum_{j=1}^{d} \int_s^t \sum_{l=1}^{N} \frac{\partial \sigma_j^i}{\partial x^l}(X(r)) D_s^q X^l(r) dB_r^j + \int_0^t \sum_{l=1}^{N} \frac{\partial b^i}{\partial x^l}(X(r)) D_s^q X^l(r) dr \tag{72}$$

for $s \leq t$ and $D_s X(t) = 0$ for $s \geq t$.

Let us prove (68). We assume for simplcity that the coefficients are bounded. Let $T < 0$ be fixed. Using Burholder's inequality one obtains for every $s \leq t \leq T$

$$\left\| D_s^q X^i(t) \right\|_p \leq C(1 + \sum_{l=1}^{N} \int_s^t \left\| D_r^q X^l(t) \right\|_p dr)$$

where C is a constant which depends on the bounds of $\sigma, \nabla\sigma$ and ∇b. Then, using Gronwall's lemma for $t \in (s, T]$ one obtains $\sup_{s \leq t \leq T} \left\| D_s^q X^i(t) \right\|_p \leq C$ where C depends on the above bounds and of T. So the proposition is proved for the first order derivatives. The proof is analogues - but much more involved from a notational point of view - for higher derivatives, so we live it out. $\square$

Remark 12 *It follows that $X_t \in Dom(L)$. The same type of arguments permit to prove that LX_t solves a SDE and*

$$\sup_{t \leq T} E \left| LX_t \right|^p \leq C_p(T) \tag{73}$$

where $C_p(T)$ depends on p, T and on the bounds of the coefficients and of their derivatives up to the order two.

The non degeneracy problem. We sketch the proof of ii).

Step 1. In a first step we will obtain the decomposition of DX and consequently for the Malliavin covariance matrix. It is well known (see [10],[11],...) that under the hypothesis (67) one may choose a version of the diffusion process X such that $x \to X_t^x$ is differentiable for every $t \geq 0$. Moreover, $Y_t =: \nabla X_t$ satisfies a SDE and, for each (t, ω), the matrix $Y_t(\omega)$ is invertible and $Z_t = Y_t^{-1}$ satisfies itself a SDE. Let us be more precise. We denote

$$Y_j^{x,i}(t) = \frac{\partial X^{x,i}(t)}{\partial x^j}.$$

Then differentiating in the SDE of X we obtain (we use the convention of summing on repeated indexes)

$$Y_j^{x,i}(t) = \delta_{ij} + \int_0^t \frac{\partial \sigma_l^i}{\partial x^k}(X_s^x) Y_j^{x,k}(s) dB_s^l + \int_0^t \frac{\partial b^i}{\partial x^k}(X_s^x) Y_j^{x,k}(s) ds.$$

We define then Z to be the solution of the SDE

$$Z_j^{x,i}(t) = \delta_{ij} - \int_0^t Z_k^{x,i} \frac{\partial \sigma_l^k}{\partial x^j}(X_s^x)(s) dB_s^l - \int_0^t Z_k^{x,i} \Big(\frac{\partial b^k}{\partial x^j} - \frac{\partial \sigma_l^k}{\partial x^r}\frac{\partial \sigma_l^r}{\partial x^j}\Big)(X_s^x)(s) ds.$$

Using Ito's formula one easily check that $dY(t)Z(t) = dZ(t)Y(t) = 0$ so that $Zt)Y(t) = Z(0)Y(0) = I$ and $Y(t)Z(t) = Y(0)Z(0) = I$. This proves that $Z = Y^{-1}$.

Using the uniqueness of the solution of the SDE (72) one obtains

$$D_s X(t) = Y_t Z_s \sigma(X_s). \tag{74}$$

and consequently

$$\gamma_{X_t} =< DX_t, DX_t >= Y_t \overline{\gamma}_t Y_t^*$$

where Y_t^* is the transposed of the matrix Y_t and $\overline{\gamma}_t$ is defined by

$$\overline{\gamma}_t = \int_0^t Z_s \sigma\sigma^*(X_s) Z_s^* ds \tag{75}$$

Using the SDE satisfied by Z standard arguments (Burkholder's inequality and Gronwall's lemma) yield $\|Z_t\|_p < \infty$ for every $p > 1$ and consequently $\|\det Z_t\|_p < \infty$. Since $Z_t = Y_t^{-1}$ it follows that $(\det Y_t)^{-1} = \det Z_t$ and consequently $\left\|(\det Y_t)^{-1}\right\|_p < \infty$ for every $p > 1$. It follows that $\left\|(\det \gamma_{X_t})^{-1}\right\|_p \leq C_p \left\|(\det \overline{\gamma}_t)^{-1}\right\|_p$ so that the hypothesis (H_{X_t}) is a consequence of

$$(H'_{X_t}) \quad \left\|\overline{\gamma}_t^{-1}\right\|_p < \infty, \quad p > 1.$$

Step 2. We are now in the frame of the previous section and so our aim is to develop $Z_t\sigma(X_t)$ in stochastic Taylor series. Using Ito's formula (see [15] Section 2.3 for the detailed computation) one obtains for every $\phi \in C^2(R^N, R^N)$

$$\begin{aligned} Z_j^i(t)\phi^j(X_t) &= \delta_{ij}\phi^j(x) + \int_0^t Z_k^i(s)[\sigma_l, \phi]^k(X(s))dB_s^l + \int_0^t Z_k^i(s)([\sigma_0, \phi]^k \\ &+ \frac{1}{2}\sum_{l=1}^d [\sigma_l, [\sigma_l, \phi]]^k)(X(s)))ds \end{aligned} \tag{76}$$

where $\sigma_l : R^N \to R^N$ is the $l'th$ column of the matrix σ for $l = 1, ..., d$ and $\sigma_0 = b - \frac{1}{2}\sigma\nabla\sigma$ (this is the drift coefficient which appears in the SDE of X when writing this equation in terms of Stratonovich integrals). We used in the above expression the convention of summation over repeated high and low indexes. Let us denote

$$\begin{aligned} T_l(\phi) &= [\sigma_l, \phi], \quad l = 1, ..., d, \\ T_0(\phi) &= [\sigma_0, \phi] + \frac{1}{2}\sum_{l=1}^d [\sigma_l, [\sigma_l, \phi]]. \end{aligned}$$

In order to get unitary notation we denote $B^0(t) = dt$ so that $dB_t^0 = dt$. We write now (76) in matrix notation

$$Z^x(t)\phi(X^x(t)) = \phi(x) + \int_0^t Z^x(s)T_l(\phi)(X^x(s))dB_s^l. \tag{77}$$

This is the basic relation (the sum is over $l = 0, ..., d$ now). We use this relation in order to obtain a development in stochastic series around $t = 0$. The first step is

$$\begin{aligned} Z^x(t)\phi(X^x(t)) &= \phi(x) + T_l(\phi)(x)B_t^l + \int_0^t (Z^x(s)T_l(\phi)(X^x(s)) - Z^x(0)T_l(\phi)(X^x(0))dB_s^l \\ &= \phi(x) + T_l(\phi)(x)B_t^l + \int_0^t \int_0^s Z^x(r)T_qT_l(\phi)(X^x(r))dB_r^q dB_s^l. \end{aligned}$$

For a multi-index $\alpha = (\alpha_1, ..., \alpha_m)$ we denote $T_\alpha = T_{\alpha_m} \circ ... T_{\alpha_1}$ Using the above procedure we obtain for each $m \in N$

$$Z^x(t)\phi(X^x(t)) = \sum_{p(\alpha)\leq m/2} T_\alpha(\phi)(x)I_\alpha(1)(t) + \sum_{p(\alpha)=(m+1)/2} I_\alpha(Z^xT_\alpha(\phi)(X^x))(t)$$

We define

$$Q_m^x(\xi) =: \inf_{|\xi|=1} \sum_{l=1}^{d} \sum_{0 \le p(\alpha) \le m/2} < T_\alpha(\sigma_l)(x), \xi >^2 .$$

Note that (H_m^x) is equivalent to $Q_m^x \neq 0$. So, using Propositoin 18 we obtain $\left\| (\det \overline{\gamma}_y)^{-1} \right\|_p < \infty$ and so (H_{X_t}) holds true. But Q_m^x contains a quantitative information: for example, in the case $m = 0, Q_0^x$ is the smaller proper value of the matrix $\sigma\sigma^*(x)$ and so $\det \sigma\sigma^*(x) \ge (Q_0^x)^N$. This permits to obtain the evaluations in (69) and (70) . We do not give here the proof - see $[K.S, 2]$.

Step 3. At this stage we apply theorem 14 and we have a smooth density $p_t(x_0, y)$ and moreover, we have polynomial bounds for the density and for its derivatives. It remains to obtain the exponential bounds under the boundedness assumptions on σ and b. The idea is the same as in section 1. Let $y \in R^N$ such that $y^i > 0, i = 1, ..., N$. One employes the integration by parts formula and Schwarz's inequality and obtains

$$p_t(x_0, y) = E(1_{I_y}(X_t^{x_0}) H_{(1,\dots,1)}(X_t^{x_0}, 1)) \le \sqrt{P(X_t^{x_0} \in I_y)} (E \left| H_{(1,\dots,1)}(X_t^{x_0}, 1) \right|^2)^{1/2}$$

where $I_y(x) = \prod_{i=1}^N [y^i, \infty)$. Using (68), (69) and the reccurcive definition of $H_{(1,\dots,1)}$ one may check that

$$(E \left| H_{(1,\dots,1)}(X_t^{x_0}, 1) \right|^2)^{1/2} \le \frac{C(1 + |x|)^q}{t^{p/2}}.$$

Moreover, using Holder'ss inequality

$$P(X_t^{x_0} \in I_y) \le \prod_{i=1}^N P(X_t^{i,x_0} \ge y^i)^{1/2^N}.$$

The evaluation of the above probabilities is standard. Recall that b is bounded. Suppose that $y^i - x_0^i > 2t \|b\|_\infty$. Then $\int_0^t b(X_s^{x_0}) ds \le \frac{y^i - x_0^i}{2}$ and so

$$P(X_t^{i,x_0} \ge y^i) \le P(M_t^{x_0} \ge \frac{y^i - x_0^i}{2})$$

where $M_t^{x_0} = \sum_{j=1}^d \int_0^t \sigma_j^i(X_s^{x_0}) dB_s^j$. Since $M_t^{x_0}$ is a continuous square integrable martingale with compensator $< M^{x_0} > (t) = \int_0^t \sum_{j=1}^d \left| \sigma_j^i(X_s^{x_0}) \right|^2 ds$ one has $M_t^{x_0} = b(< M^{x_0} > (t))$ where b is a standard Brownian motion.

Note that $< M^{x_0} > (t) \le d \times t \times \|\sigma\|_\infty^2$ so that $|M_t^{x_0}| \le \sup_{s \le d \times t \times \|\sigma\|_\infty^2} |b(s)|$.

Finally using Doob's inequality and elementary evaluations for gaussian random variables we obtain

$$
\begin{aligned}
P(X_t^{i,x_0} &\geq y^i) \leq P(M_t^{x_0} \geq \frac{y^i - x_0^i}{2}) \leq P(\sup_{s \leq d \times t \times \|\sigma\|_\infty^2} |b(s)| \geq \frac{y^i - x_0^i}{2}) \\
&\leq 4P(\left| b(d \times t \times \|\sigma\|_\infty^2) \right| \geq \frac{y^i - x_0^i}{2}) \leq \frac{C}{\sqrt{t}} \exp(C' \frac{(y^i - x_0^i)^2}{2t}).
\end{aligned}
$$

So we proved the exponential bound, at list if $y^i \to +\infty, i = 1, ..., N$. Suppose now that $y^i \to +\infty, i = 1, ..., N-1$ and $y^N \to -\infty$. Then one has to replace I_y by $\prod_{i=1}^{N-1} [y^i, \infty) \times (-\infty, y^N]$ in the integration by parts formula. In such a way one obtains the exponential bounds for every y such that $|y| \to \infty$. In order to prove the bounds for the derivatives one has to employ integration by parts formulas of higher order, as in the proof of Lemma 2 in section 1. $\square$

References

[1] V. Bally: On the connection between the Malliavin covariance matrix and Hormander's condition. J. Functional Anal. 96 (1991) 219-255.

[2] D.R. Bell: The Maliiavin Calculus. Pitman Monographs and Surveys in Pure and Applied Math. 34, Longman and Wiley, 1987.

[3] G. Ben Arous: Flots et séries de Taylor stochastiques; PTRF 81(1989), 29-77

[4] K. Bichteler, J.B. Gravereaux and J. Jacod: Malliavin Calculus for Processes with Jumps, Stochastic Monographs Vol. 2, Gordon and Breach Publ., 1987.

[5] N. Bouleau and F. Hirsch: Dirichlet Forms and Analysis on Wiener Space. de Gruyter Studies in math. 14, Walter de Gruyter, 1991

[6] G. Da Prato, P. Malliavin, D. Nuamart: Compact Families of Wiener Functionals, CRAS tome 315, série 1, pg 1287-1291, année 1992.

[7] M. Hitsuda: Formula for Brownian motion derivatives. Publ. Fac. of Integrated Arts and Sciences Hiroshima Univ. 3 (1979) 1-15.

[8] E. Fournier, J.M. Lasry, J. Lebouchoux, P-L Lions, N. Touzi: Applications of Malliavin Calculus to Monte Carlo methods in finance. Finance Stochast. 3, 391-412 (1999)

[9] E. Fournier, J.M. Lasry, J. Lebouchoux, P-L Lions: Applications of Malliavin Calculus to Monte Carlo methods in finance II. Finance Stochast. 5, 201-236 (2001).

[10] N. Ikeda and S. Watanabe : Stochastic differential equations and diffusion processes, Second Edition, North-Holland 1989.

[11] H. Kunita: Stochastic flows and Stochastic Differential Equations, Cambridge Univ. Press, 1988.

[12] S. Kusoucka and D. W. Stroock: Applications of the Malliavin Calculus II. J. Fac. Sci. Univ. Tokyo Sect IA Math. 32 (1985) 1-76.

[13] P. Malliavin: Stochastic calculus of variations and hypoelliptic operators.. In Proc. Inter. Symp. on Stoch. Diff. Equations, Kyoto 1976, Wily 1978, 195-263.

[14] J. Norris: Simplified Malliavin Calculus. In Seminaire de Probabilités XX, lecture Notes in math. 1204 (1986) 101-130.

[15] D. Nualart The Malliavin Calculus and Related Topics.In probability and its Applications, Springer-Verlag, 1995.

[16] D. Ocone: Malliavin calculus and stochastic integral representation of diffusion processes. Stochastics 12 (1984) 161-185.

[17] I. Sigekawa: Derivatives of Weiner functionals and absolute continuity of induced measures. J. Math. Kyoto Univ. 20-2(1980) 263-289.

[18] A. V. Skorohod: On a generalization of a stochastic integral. Thory Probab. Appl. 20 (1975) 219-233.

[19] F.J. Vasquez- Abad: A course on sensitivity analysis for gradient estimation of DES performance measures.

[20] S. Watanabe: Lectures on Stochastic Differential Equations and Malliavin calculus, Tata Institute of Fundamental Research, Springer-Verlag, 1984.

[21] D. Williams: To begin at the beginning.. In Stochastic Integrals, Lecture notes in Math. 851 (1981) 1-55.

Contents

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399